UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE FÍSICA



# Image Reconstruction Algorithm Implementation for the easyPET: a didactic and pre-clinical PET system

Pedro Miguel Martins de Sousa e Sá

**Mestrado Integrado em Engenharia Biomédica e Biofísica**

Perfil em Radiações em Diagnóstico e Terapia

Dissertação orientada por:

Dr. Nuno Matela

2017

# ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the people that are part of the easyPET project and with whom I have worked with. To Prof. Dr. João Veloso, Prof. Dr. Pedro Almeida, and Prof. Dr. Nuno Matela for having allowed me to participate in such innovative project and develop the work presented in this thesis. To Pedro Correia, the right person to work with, given his friendliness, and willingness and objectivity when discussing matters. To Joana Menoita who has produced part of the results presented here. Finally, to all the participants in this project, whom have brought the easyPET so far, might it go as far as intended.

To my supervisor, Prof. Dr. Nuno Matela, with whom I deeply appreciate having worked with, given his friendliness and availability, with the occasional football discussions and moderate club rivalry. Even more, I would like to acknowledge the interest and knowledge on Biomedical Engineering he has conveyed to me during my academical journey, culminating with this thesis. I truly am thankful to him.

To all the persons I have met at IBEB until this moment, all the Professors, researchers, and students. People have come and go, but it remains a special place to work in, and I could not be more thankful. On a special notice, I would like to acknowledge Beatriz Lampreia, for being a friendly face at all times. For the time I have spent at IBEB, most was in the company of Nuno Santos, with whom I am very happy to have become friends with. Currently, me and Leon Vink are still resilient in leaving.

To the people I have met throughout my academic path, in one way or another, you have contributed to my success. To Francisco Geraldes, Joaquim Costa, and Rui Marcelino, it has been a good ride.

Finally, and after all, it was up to my family to put me in this world. The effort and dedication I have seen and know in my parents and grandparents have no bounds. Life does not always come easy, but it is no excuse to stop and not work for the best. To my brother, well, being the youngest is no longer terrible. I sure am glad you did not push me off the nest when you had the chance.

# RESUMO

Tomografia por Emissão de Positrões (PET) é uma técnica de imagiologia funcional, utilizada para observar processos biológicos. O conceito de tomografia por emissão foi introduzido durante a década de 1950, sendo que foi apenas com o desenvolvimento de radiofármacos na década de 1970, que esta técnica começou a ser utilizada em medicina. Nos últimos 20 anos, o avanço tecnológico tornou os sistemas PET numa ferramenta altamente qualificada para imagiologia funcional. Neste período, o aparecimento de sistemas PET-CT veio colmatar as deficiências produzidas pela PET ao nível de imagem estrutural, com a combinação desta técnica funcional com a de Tomografia Computadorizada (CT). A evolução da tecnologia PET foi também acompanhada pela evolução da tecnologia para produção de radiofármacos, incluindo os radionuclídeos, bem como do conhecimento médico relativo aos processos biológicos humanos. Aliando esta tecnologia e conhecimento, tornou-se possível traçar moléculas com funções metabólicas nos diversos sistemas do corpo humano e, assim, produzir uma variedade de imagens funcionais.

Dado o tipo de imagem produzida pela técnica PET, é bastante comum associar-lhe o diagnóstico de doenças cancerígenas, cuja principal característica é a desregulação metabólica celular no organismo. Tendo em vista o aumento esperado da incidência de cancro em Portugal e na Europa, tendo já sido atingida uma incidência nacional, em 2010, de 444,50 pessoas em cada 100.000 (números avançados pela DGS, 2015), a utilização de técnicas que permitam o diagnóstico precoce destas doenças é de elevada importância. Posto isto, e apesar do constante crescimento do gasto público em cuidados médicos relativos ao diagnóstico e tratamento de cancro, estão a ser postos cada vez mais esforços e fundos para que o processo de Investigação e Desenvolvimento (I&D) relacionado com esta doença seja célere. São constantemente desenvolvidas novas e melhores técnicas de imagiologia, que permitem diagnósticos mais precoces e precisos, enquanto ajudam na aplicação de planos de tratamento mais eficazes que, consequentemente, levam a um gasto público mais eficiente.

Os sistemas PET inserem-se neste contexto e, uma vez permitindo imagem altamente sensível a processos funcionais, facilmente se generalizaram no meio médico e académico. Os sistemas direcionados a aplicações relacionadas com a medicina humana têm como função observar processos biológicos, com a finalidade de um diagnóstico médico ou estudo. Sistemas pré-clínicos, direcionados a estudos com animais pequenos, têm o propósito de auxiliar a investigação relacionada com os estudos preliminares de doenças que afetem o ser humano. Finalmente, e sendo o grupo com menor oferta comercial, os sistemas PET didáticos possibilitam uma melhor formação de pessoal responsável pelo futuro uso e I&D relacionados com esta tecnologia. No entanto, a tecnologia utilizada nestes três tipos de sistemas encarece consideravelmente o seu valor comercial sendo que, contrariamente ao que seria de esperar, os preços dos sistemas pré-clínicos não se diferenciam consideravelmente dos sistemas para humanos. O encarecimento destes sistemas deve-se ao facto de que toda a tecnologia a eles associada tem características mais dispendiosas de produzir. No caso dos sistemas didáticos, simplesmente não existe o incentivo necessário à sua produção e compra.

 É neste contexto que surge o easyPET. O design inovador, constituído por apenas duas colunas de detetores opostos, e tirando partido de uma atuação sobre dois eixos de rotação, faz deste sistema ideal para entrar no mercado em duas vertentes. A primeira, constituída apenas por um detetor em cada coluna, está destinada a ter um papel didático. A segunda, tirando partido de colunas com múltiplos detetores, foi desenhada para entrar no mercado de sistemas pré-clínicos. Em ambos os casos, a principal característica do easyPET, e a que o destaca dos restantes sistemas, é o seu reduzido número de detetores, que resulta num reduzido custo de produção. Através da implementação de um número reduzido de detetores e, consequentemente, reduzida eletrónica, é possível obter um custo final da máquina inferior.

No entanto, é sempre necessário garantir que os dados obtidos em tal sistema correspondam a imagens com as características necessárias, sendo que o processo de reconstrução de imagem é bastante importante.

O trabalho apresentado nesta tese tem como objetivo a implementação de um método de reconstrução de imagem a duas dimensões, dedicado ao sistema easyPET. Para tal, foi considerado um algoritmo estatístico iterativo que se baseia na Maximização da Estimativa da Máxima Verosimilhança (ML-EM), introduzido por Shepp e Vardi em 1982. Desde então, tem sido largamente explorado e, inclusive, dando aso a outras versões bastante comuns em reconstrução de imagem PET, como é caso da Maximização da Espectativa usando Subgrupos Ordenados (OS-EM).

A implementação do algoritmo escolhido foi feita no software Matlab. Para computar a unidade básica do algoritmo, a Linha de Resposta (LOR), foi implementado o método ray-driven. Por forma a otimizar a construção da matriz de sistema utilizada neste algoritmo, foram implementadas simetrias de geometria. Esta otimização baseou-se na consideração de que a geometria do sistema easyPET pode ser dividida em quadrantes, sendo que um único quadrante consegue descrever os restantes três. Além disso, foram também implementadas otimizações ao nível estrutural do código escrito em Matlab. Estas foram feitas tendo em conta o aumento na facilidade de acesso à memória através da utilização variáveis para rápido indexamento. Foram também implementados dois métodos de regularização de dados: filtragem gaussiana entre iterações e um *root prior* baseado na mediana. Por forma a comparar, mais tarde, os resultados obtidos através do algoritmo implementado, foi também implementado o método de reconstrução de Retroprojeção Filtrada (FBP). Por último, foi implementada uma interface para o utilizador, utilizando a aplicação GUIDE do Matlab. Esta interface tem como objetivo servir de ponte entre o sistema didático easyPET e o utilizador, para que a experiência de utilização seja otimizada.

Por forma a delinear o teste ao sistema easyPET e ao algoritmo ML-EM implementado, foram seguidas as normas NEMA. Este é um conjunto de normas que tem como objetivo padronizar a análise realizada a sistemas de imagem médica. Para tal, foram adquiridos e simulados ficheiros de dados com uma fonte pontual a 5, 10, 15 e 25 mm do centro do campo de visão do sistema (FOV) e utilizando um par de detetores com 2x2x30 mm$^3$. Para realizar a análise de resultados, os dados foram reconstruídos utilizando a FBP implementada, e foi medida a FWHM e FWTM da fonte reconstruída. O mesmo procedimento foi aplicado, mas reconstruindo os dados através do algoritmo ML-EM, utilizando o filtro gaussiano, o MRP, e não utilizando qualquer método de regularização de dados (nativo). Por forma a comparar os métodos de regularização de dados, foi também realizada uma medição do rácio sinal-ruído (SNR). Os resultados foram obtidos para imagens reconstruídas com um pixel de, aproximadamente, 0.25x0.25 mm$^2$, correspondendo a imagens de 230x230 pixeis.

Os primeiros resultados foram obtidos a fim de determinar qual a iteração em que se começaria a observar a estabilização das imagens reconstruídas. Para algoritmo ML-EM implementado e o tipo de dados utilizados, foi observado que a partir da 10ª iteração o algoritmo ML-EM converge. Através das medidas para a FWHM e FWTM observou-se, também, que os dados obtidos experimentalmente se diferenciam dos resultados obtidos sobre os dados simulados. Isto levou a que, fora dos objetivos deste trabalho, fossem realizados mais testes utilizando dados experimentais e, que daqui em diante, apenas fossem utilizados dados obtidos através de simulação Monte Carlo, por razões de conveniência na precisão da colocação da fonte pontual. De seguida, comparam-se os dados obtidos através da FBP e o algoritmo ML-EM nativo. Para o primeiro caso foram medidas FWHM de 1.5x1.5 mm$^2$, enquanto que para o segundo foram atingidos valores de 1.2x1.2 mm$^2$. Para os métodos de regularização de dados foram medidos valores de resolução semelhantes ou inferiores, sendo que estes resultaram num aumento da qualidade da reconstrução da fonte, observado através do aumento no valor de SNR medido.

O trabalho apresentado nesta tese revela, não só a validação do algoritmo de reconstrução proposto, mas também o bom funcionamento e potencialidades do sistema easyPET. Pelos resultados obtidos através das normas NEMA, é possível observar que este sistema vai ao encontro do estado de arte. Mais ainda, através de um método de reconstrução dedicado ao easyPET é possível otimizar os resultados obtidos. Com o avançar do projeto no qual este trabalho esteve inserido, é de esperar que o modelo a três dimensões pré-clínico easyPET irá produzir melhores resultados. De frisar que o sistema easyPET didático se encontra na sua fase final e que os resultados obtidos são bastante satisfatórios tendo em conta a finalidade deste sistema.


**Palavras Chave:**    Tomografia por Emissão de Positrões (PET)

Reconstrução de Imagem em PET

Maximização da Espectativa da Máxima Verosimilhança (ML-EM)

Resolução Espacial

Normas NEMA

# ABSTRACT

The easyPET scanner has an innovative design, comprising only two array columns facing each other, and with an actuation defined by two rotation axes. Using this design, two approaches have been taken. The first concerns to a didactic PET scanner, where the arrays of detectors are comprised of only one detector each, and it is meant to be a simple 2-dimensional PET scanner for educational purposes. The second corresponds to a pre-clinical scanner, with the arrays having multiple detectors, meant to acquire 3-dimensional data. Given the geometry of the system, there is no concern with the effects of not measuring the Depth-of-Interaction (DOI), and a resolution of 1-1.5 mm is expected with the didactic system, improving with the pre-clinical. The work presented in this thesis deals with 2D image reconstruction for the easyPET scanners.

The unconventional nature of the acquisition geometry, the large amount of data to be processed, the complexity of implementing a PET image reconstruction algorithm, and the implementation of data regularization methods, gaussian filtering and Median Root Prior (MRP), were addressed in this thesis. For this, the Matlab software was used to implement the ML-EM algorithm. Alongside, several optimizations were also implemented in order to convey a better computational performance to the algorithm. These optimizations refer to using geometry symmetries and fast indexing approaches. Moreover, a user interface was created so as to enhance the user experience for the didactic easyPET system.

The validation of the implemented algorithm was performed using Monte Carlo simulated, and acquired data. The first results obtained indicate that the optimizations implemented on the algorithm have successfully reduced the image reconstruction time. On top of that, the system was tested according to the NEMA rules. A comparison was then made between reconstructed images produced by using Filtered Back Projection (FBP), the native ML-EM implementation, the ML-EM algorithm using inter-iteration gaussian filtering, and the ML-EM algorithm implemented with the MRP. This comparison was made through the calculation of FWHM, FWTM, and SNR, at different spatial positions. The results obtained reveal an approximate $1.5x\ 1.5\ mm^2$ FWHM source resolution in the FOV, when recurring to FBP, and $1.2x\ 1.2\ mm^2$ for the native ML-EM algorithm. The implemented data regularization methods produced similar or improved spatial resolution results, whilst improving the source's SNR values. The results obtained show the potential in the easyPET systems. Since the didactic scanner is already on its final stage, the next step will be to further test the pre-clinical system.

# INDEX

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **2D** | Two dimensional |
| **3D** | Three dimensional |
| **AOR** | Area of Response |
| **BSO** | Bismuth Silicone Oxide |
| **CPU** | Central Processing Unit |
| **CT** | Computed Tomography |
| **DOI** | Depth of Interaction |
| **EM** | Expectation Maximization |
| **FBP** | Filtered Back Projection |
| **FDG** | Fludeoxyglucose |
| **FOV** | Field of View |
| **FWHM** | Full Width at Half Maximum |
| **FWTM** | Full Width at Tenth Maximum |
| **GSO** | Gadolinium Oxyorthosilicate |
| **GPU** | Graphics Processing Unit |
| **GUI** | Graphical User Interface |
| **GUIDE** | Graphical User Interface Development Environment |
| **keV** | Kilo Electronvolt |
| **LOR** | Line of Response |
| **LYSO** | Lutetium-yttrium Oxyorthosilicate |
| **ML** | Maximum Likelihood |
| **ML-EM** | Maximum Likelihood Expectation Maximization |
| **MR** | Magnetic Resonance |
| **MRP** | Median Root Prior |
| **OS-EM** | Ordered Subset Expectation Maximization |
| **PET** | Positron Emission Tomography |
| **RAM** | Random Access Memory |
| **ROI** | Region of Interest |
| **SNR** | Signal-to-noise Ratio |
| **SPECT** | Single-photon Emission Computed Tomography |
| **TOF** | Time of Flight |
| **TOR** | Tube of Response |

# 1 Introduction

## 1.1 PET Technique Evolution and Role in Medicine

Positron Emission Tomography (PET) is a nuclear functional imaging technique, used to observe biological processes. The concept of emission tomography was first introduced in the 1950s. Soon after, the first PET scanner appeared. Yet, it was only in the 1970s, with the development of radiopharmaceuticals, that PET imaging technique saw its first breakthrough into medicine. The new millennium saw its acceptance in the medical community widened with the introduction of PET-CT imaging systems, combining both functional and structural imaging techniques.

In medicine, PET technique helps practitioners study and diagnose diseases that carry specific biomarkers, or have localized high concentrations of specific biomolecules. Most commonly, PET technique is associated with cancer diagnosis, as the abnormal cellular growth relates to erratic metabolic levels.

In 2016, and only in the USA, an estimated 1 685 200 new cases of cancer were diagnosed, while 595 690 have died. This represents an incidence of a staggering 454.8 per 100 000 persons, and a mortality of 171.2 per 100 000 [1]. Although cancer incidence is increasing, frequently associated with constant and increasing exposure to risk factors, the overall death rate for this disease is decreasing. With an ever-rising medical care expenditure on cancer diagnosis and treatment, funds and efforts are being put to cancer related research and development (R&D). New and enhanced imaging techniques are allowing more precise and earlier diagnosis, while helping form more effective treatment plans, leading to a more efficient funds expenditure.

Research on PET appears in this context. Being a highly sensitive functional imaging technique, it has become widely spread, and current PET systems range from clinical to educational applications. Systems for human applications have a purpose in diagnosing and studying biological processes. Devices for small animal's studies are meant to aid researchers in better study diseases which relate to the human being. Finally, and being the scarcer group, didactic PET systems are helping better train people who will perform PET related work.

However, there are two main drawbacks with this technique. The first being the infrastructure needed to provide radionuclides to perform a PET scan. The second, and the one that relates the most with medical care and overall R&D, is total expenditure. Disregarding all the funds needed to start a R&D project, imaging systems, in this case PET, build up to be one of the most expensive equipment in any medical care or research facility. One would guess that small animals and didactic devices are cheaper whilst comparing to human systems. Yet, small animal's devices use more expensive technology that allows smaller image resolution. As for didactic systems, even though they are cheaper by only allowing reduced image resolution and quality, academic institutions are not able to cope with the prices.

The following written work relates to PET image reconstruction. The project in which it was developed aims to tackle the need for two different devices: a didactic and a small animal's system, both being commercially competitive, specs wise, and as affordable as possible.

# 2 Background

In this chapter, an introduction will be made on all relevant topics regarding PET technique. Naturally, some topics will be kept as simple as possible, as it is out of this work's scope to thoroughly explain all topics.

## 2.1 Molecules and Tracers

In order to study a disease using the PET technique, one has to arrange a way to track down the metabolic processes behind it. For this, two things are required: a molecule that acts as a reactant in the metabolic process to observe, and a radionuclide that can be used as a tracer for the molecule used. With these, it is possible to create a radiopharmaceutical that, ultimately, will be observed in a PET scan.

Metabolic processes occur all throughout the human body. However, when comparing with healthy subjects, diseases create changes in metabolic patterns. Depending on the focused metabolic process to be observed, reactant molecules must be chosen accordingly. Even though PET is often associated with cancer disease, where the abnormal cellular growth is known to lead to the increase region intake of glucose, radiopharmaceuticals are produced to best fit the object disease. As can be seen in Table 2.1, diseases range from stroke to lung perfusion, bone cancer, and organ failure.

What transmits the location of such molecules are radioactive tracers which, when combined with biological reactants, creates a radiopharmaceutical. These tracers are chosen depending on their half-life, which will be discussed next, commercial availability, and affinity to the chosen molecule. Some of the most common radiopharmaceuticals and their purposes can be seen in Table 2.1.

**Table 2.1:** Common radionuclides used in PET, their half-lives, correspondent rediopharmaceutic, and intended object of study.

| RADIONUCLIDE | HALF-LIFE (MIN) | RADIOPHARMACEUTICAL | TARGET |
|---|---|---|---|
| **CARBON-11** | 20.3 | C11-PABA | Pancreatic studies |
| **FLUORINE-18** | 109 | F18-Flurodeoxyglucose (FDG) | Tumor and myocardial imaging |
| | | F18-Sodium fluorine | Bone imaging |
| **GALIUM-68** | 68 | Ga68-Dotatate | Neuroendocrine tumor imaging |
| **OXYGEN-15** | 2 | O15-$H_2$O | Cerebral Blood flow imaging |

## 2.2 Basic Physics

### 2.2.1 β+ decay

As the name states, this imaging technique relies on positron emission. Used positron emitting radionuclides are produced in a ciclotron and, usually, the isotopes are chosen to have short half-lives, when comparing to other well-known radioactive isotopes (Uranium, Plutonium, etc). The shorter half-lives are compatible with the biological processes aimed for observation. Stable nuclei like oxygen-16, carbon-14, etc., found in metabolic reactants, such as water, glucose, and ammonia, are replaced by the correspondent isotopes (e.g. oxygen-15, carbon-11, etc.). Yet, radionuclides can also be chosen to trace synthetic drugs delivered to our body. All around, these labelled compounds are called radiotracers and

are used to find the source or distribution of a given disease in PET exams. Note that isotopes are also chosen based on their half-lives. For optimal results, the time any given compound takes to reach its intended target should match, within reason, with the isotope's half-life [2].

Positron emission ($\beta+$ decay) is a subtype of radioactive decay. It is represented by a proton turning into a neutron by emitting a positron (Equation 2.1).

$$p \rightarrow e^+ + n + v_e \tag{2.1}$$

Positron emission takes place when a nucleus is unstable due to an imbalance between protons and neutrons. In $\beta+$ decay, this state is balanced by a proton emitting a positron according to Equation 2.1. A nucleus goes into decay so as to lose energy and become more stable [3]. The results of the decay are the emission of a positron, neutrino, and neutron. In nuclei decay, as seen in Equation 2.2, the daughter nucleus represents the neutron in Equation 2.1.

$$^{23}_{12}Mg \rightarrow {}^{23}_{11}Na + e^+ + v_e \tag{2.2}$$

## 2.2.2 Positron Annihilation

When a positron is emitted, it collides with any surrounding electron. This collision results in the annihilation of both particles and the forming of a pair of gamma ray photons [3], as seen in Equation 2.3.

$$e^- + e^+ \rightarrow \gamma + \gamma \tag{2.3}$$
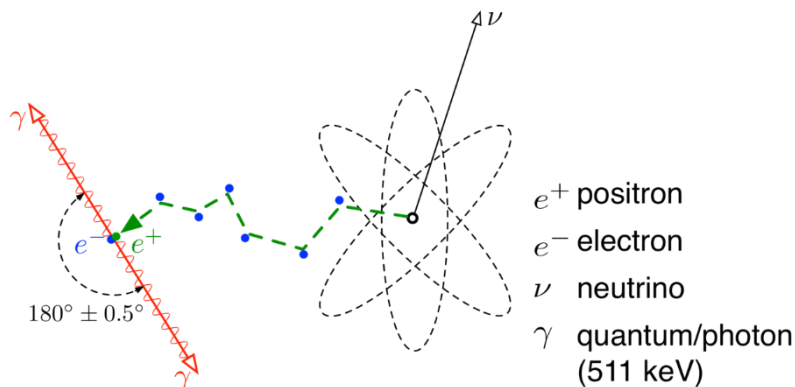


**Figure 2.1:** Annihilation process known in elementary physics. A positron (e+) is emitted from the atomic nucleus together with a neutrino (v). The positron is ejected randomly and travels through matter until colliding with an electron (e-), hence producing an annihilation [4].

The photon pair abides to, among several others, conservation of linear momentum and total energy. Much so that, it is considered that each photon has an energy of 511 keV and both travel in opposite directions (see Figure 2.1). After the forming of the pair, each photon travels in the direction of the system's detectors in which they are detected, and information is processed to later allow image reconstruction.

The assumption that the photons have the aforementioned characteristics only arises for simplicity reasons. The 511 keV energy value is only true if the positron-electron pair has a zero relative velocity. As for travelling in opposite directions, the 180° value refers to an approximation over the conservation of linear momentum, since it is considered that both the positron and electron have low kinetic energy. Yet, as the photons are not 180° apart and since it is not possible to exactly measure the difference, there will always be a degree of uncertainty affecting the signal-to-noise ratio of the data.

## 2.3  Event Detection, Coincidence, and Detector Bins

For a detection to be processed, the photon must be converted into electric current that allows its registry. Schematically, this conversion process can be made by means of a scintillator followed by a photodetector (see Figure 2.2).
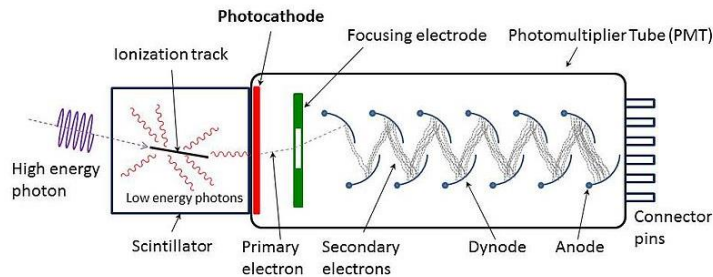


**Figure 2.2:** Schematic of a photomultiplier tube coupled to a scintillator [5].

### 2.3.1  Scintillator

When the photon interacts with the scintillator material it produces photons mostly belonging to the light spectrum, ranging from UV to infra-red radiation. The scintillator's sensitivity is also improved by its material. Heavy ions used in the scintillator lattice allow higher stopping powers, and a more compact detector. The higher stopping power results in a better segmentation between $\gamma$-rays, reducing the range of Compton scattered photons, improving the detector's spatial resolution. This also results in an increased scintillator's photo-fraction [6]. High operation speed is also desired. For precise time measurement, the scintillator needs to have short rise and decay times, which will optimize coincidence detection, enabling time of flight capabilities, and decrease the dead-time. Common crystals used in scintillators meant for PET detectors, and their properties, are depicted in Table 2.2.

**Table 2.2:** Designations of common crystals used in PET detector's scintillators, and their main attributes, having NaI as reference [7].

| CRYSTAL | PARTICULARITIES |
|---------|-----------------|
| BGO | High stopping power and low optical yield |
| GSO | Good energy resolution |
| LYSO | High optical yield and energy resolution |
| NAI | Highest optical yield and lowest stopping power. |

### 2.3.2  Photodetector

When the light reaches the photodetector, multiple electrons are emitted due to photoelectric effect, resulting in an electric current. Much like the scintillators, photodetectors help improve the intrinsic total efficiency of a detector. The main difference with these is that they operate in different wavelengths and are designed to produced electric current. In PET, photodetectors are most commonly either photomultiplier tubes (PMT) or semiconductor based photodiodes.

PMTs have high gain in photoelectric conversion, resulting in high signal-to-noise ratios. However, these have a low ratio between the incident photons and the primary produced electrons, and have a considerable size when comparing with its field of view, which is not desirable in highly dedicated scanners. More advanced PMTs, called PS-PMT, are currently being used to design high resolution PET scanners.

On the other hand, semiconductor photodetector arrays offer a greater variety of capabilities, and the most used in PET is the avalanche photodiode (APD). Semiconductor photodetectors are more compact, offer a higher efficiency, and can be used in PET-MR scanner, due to not being sensitive to magnetic fields. The most common substrate for these is Silicon, hence resulting in the widely known Silicon Photomultipliers (SiPMs), which are comprised of a series of sequentially connected Silicon APDs Even though SiPMs have become widely used, these have greatly reduced sensitivities when working at high temperatures and carry a bias voltage [8].

### 2.3.3  PET Detector and Coincidence Detection

The aforementioned hardware greatly increases the price of PET systems, with more efficient and precise components being constantly developed. Current detectors can measure the energy of the gamma photon at arrival, its arrival time with even more precision, its Depth of Interaction (DOI) with the scintillator, and have reduced cooling times, among other capabilities. These add more information to each detection, later allowing energy cut offs, Time of Flight (TOF) mode, DOI correction, or enlarged data statistics through enhanced detector efficiency. All the information obtained at each detection is processed in order for the system to output its acquired data.

Each time a detector is triggered, an entry is made in a list mode file. This entry comprises the spatial coordinates of the given detector and any other relevant information on the detection. Two detections that happen at the same time in a possible pair of detectors are known as a coincidence. A coincidence corresponds to a detector pair being triggered, with this detector pair defining a unique Line of Response (LOR), which will be introduced next. Later, when a histogram of the data is built, detector pairs are represented by bins in a way that the photons produced in annihilation events get deposited in them, making them bins for annihilation counts. Much like each detector or detector pair, a bin can also be described by specific spatial coordinates only dependent on the system's geometry.  Coincidence processing is performed when dealing with the list mode file, which will be introduced next.

# 3  PET Data Output and Representation

Since PET requires data reconstruction to obtain images, it is important to understand and specify how such data is outputted from the system, and how it can be represented. What follows is an introduction to such topics.

## 3.1  System's Output – PET Data

### 3.1.1  List Mode and its Post-Processing

After each detected photon, the system processes all the information considered vital for reconstruction and stores it as a list mode. In this way, a list mode file will have an entry for each detection times the number of variables stored per detection. This being, for N detections and M variables per coincidence, the list mode file will have NxM entries.

Concerning the variables stored, they must always carry information that locates the detector triggered by a photon. Any other information is secondary. However, with modern technology it became possible to acquire more information on photon detection. Such information, like accurately accounting the precise time the photon was detected or the energy of a detected photon, allows modern PET systems to achieve better results by using time of flight modes (TOF) or make energy cut-offs, for example. Note that, vital information varies from manufacturer, to system's geometry, and even to reconstruction methods. Spatial information can be stored as a set of spatial coordinates, a set of codes for each active detector, or even as a time coordinate [7].

The list mode file can be further processed. Usually, entries are paired according to their time stamp. Detections that have occurred within a reasonable time interval are considered to belong to the same annihilation event. Hence, two entries are substituted by a single one describing the two triggered detectors. This new single entry represents a count for a given pair of detectors. A linearly possible combination of two detectors is called a Line of Response (see Figure 3.1). As we will see Chapter 5, one can further process the list mode file in order to enhance data and future images quality.

## 3.2  LOR

A Line of Response (LOR) is the 2D representation of the line defined by a detector pair. It translates the line where a detected event has a non-zero probability of having taken place (see Figure 3.1). When a coincidence is obtained, two detectors are activated. By drawing a line between those two detectors, a LOR is obtained. The importance of the LOR is that it limits the space where one can trace a given event. Much so that, a given annihilation has maximum probability of having occurred in the LOR space. LORs can also be traced in the 3D space, be traced as an Area of Response (AOR), or even Tube of Response (TOR). In Chapter 9, a more thorough explanation will be made on how the approach to LOR calculation was made. Throughout this thesis, whenever it is read the acronym LOR, mainly in the introduction, be mindful that this is made for simplicity sake. LOR, AOR, and TOR share the same probabilistic meaning, though the three are presented with different spatial representations.
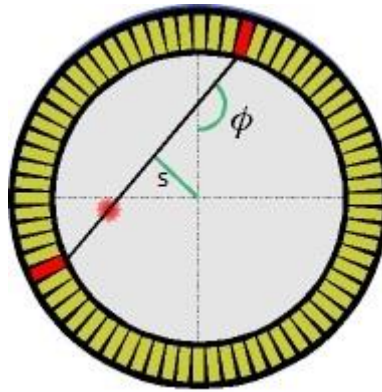
**Figure 3.1:** Representation of a LOR in the 2D space, given an annihilation that activates the detectors in red. s is the distance to the center of the FOV. $\phi$ is the angle of rotation defined by s. Adapted from [9].

### 3.2.1 LOR Computing and System Matrix

When computing, a LOR can be represented in two ways: as a line equation or as a representation in the imaged space. To obtain the second, as a computed image, which is a discretized representation of space in pixels or voxels, a LOR must too be discrete. For image reconstruction, all LORs must be used at some point. For this, we build the system matrix which is the set of pixelized LORs. The system matrix is comprised of all possible LORs a system can produce and will be presented in Chapter 9.4.2.

## 3.3 Sinogram

Sinograms are the most traditional way to represent data acquired in a PET scan. They are a histogram of the post-processed list-mode file (see Figure 3.2). The dimensions of a sinogram correspond to a distance s and an angle $\theta$. As seen in Figure 3.1, (s, $\theta$) coordinates pair describes the orientation of the LOR, where s is the tistance of a given LOR to the center of the FOV and $\theta$ is the angle between the LOR and the vertical axis. Evidently, the range of possible s's and $\theta$'s is limited by the system's geometry. Each entry in a sinogram is called a bin and has a value corresponding to the total amount of coincidences the detector pair described by the (s, $\theta$) has counted, hence being called a bin.
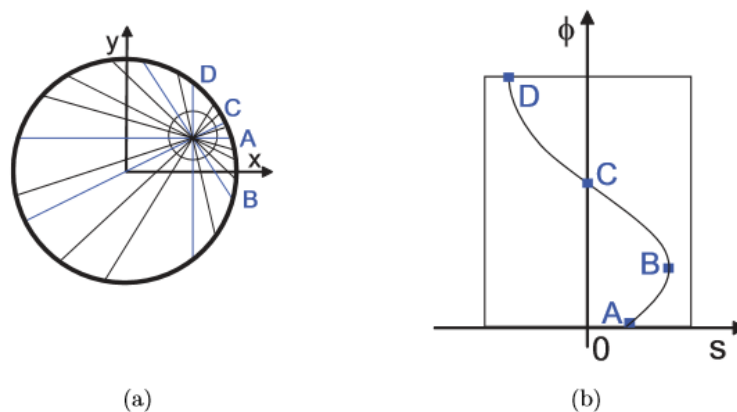


**Figure 3.2:** A centered point source and an off-centered point source in the scanner (a) describe, respectively, a straight line and a sinusoidal line in the sinogram (b). Adapted from *[10]*.

Mathematically, tracer distribution of an imaged object can be seen as a function with unknown density. A sinogram is the radon transform of that function. Imaging tracer distribution can be performed by applying the inverse Radon transform on the sinogram [11].

7

# 4 Image Reconstruction

When working with tomography techniques, image reconstruction is needed. This step takes the data acquired by the system and reconstructs it to obtain an image that best fits the object, as seen by the technique used. For instance, if one works with x-rays, the object is seen as regions that produce less or more attenuation to them. For ultrasounds, regions are seen as being more or less ultrasound reflective. However, PET differs from the later. Being an emission image technique, it relies on tracer presence throughout the internal organization of the object. As a result, the objective of PET image reconstruction is to image tracer distribution.

## 4.1 Image Reconstruction Algorithm

Image reconstruction algorithms are designed to fit the data they work with and the available computational power. At the very beginning, PET image reconstruction was based on analytical algorithms. These algorithms assume the data has little to no amount of noise, and their linear design and behavior forbids any complex image correction throughout the reconstruction process. However, with the ever-increasing computational power and acceleration techniques, iterative approaches to PET data reconstruction have appeared. Of most interest, statistic iterative algorithms, as these assume a better model for the Poisson noise distribution present in PET data, and can be shaped to incorporate noise reduction methods. The statistic algorithm exploited in this work was the Maximum Likelihood Expectation Maximization (ML-EM) algorithm. Although, we will be comparing the results obtained by this algorithm with Filtered Back Projection (FBP), for reasons that will be explained later on.

## 4.2 Iterative Algorithms and Expectation Maximization (EM)

As was previously mentioned, tracer distribution is a function with unknown density. Image reconstruction is used to obtain the cross-sectional image reflecting tracer distribution. The reconstruction algorithm of most relevance to this work is based on Expectation Maximization (EM) [12]. Through EM, tracer distribution is as follows:

$$E[p(\lambda)] = \int f(x, y, z) \qquad (4.1)$$

Where $E[p(\lambda)]$ is the probability expectation of a certain LOR, and $f(x, y, z)$ is the tracer spatial distribution.

With the improvements on computational power and capabilities, more complex image reconstruction algorithms are being implemented. Iterative algorithms have become more common, as they achieve better results. However, these expend more time, and require more computational power than analytical and recursive algorithms.

Iterative methods appear in the context of computational mathematics. These methods allow a sequence of improving approximate solutions, through successive forward and back projections on a given problem. Although an iterative algorithm may only converge to a non-absolute solution [13], it is preferable to use this approach when dealing with high complexity problems, as is the case of reconstructing PET data.

Most common algorithms for PET image reconstruction are based in the statistical Expectation Maximization (EM) approach, such as the ML-EM, Ordered Subset EM (OS-EM), or Maximum a Posteriori (MAP) [14]. An EM iteration comprises an Expectation and a Maximization phase. The flow chart for an iterative approach is depicted in Figure 4.1. In this case, the E-phase corresponds to a forward projection, where an estimated image is derived. Through the M-phase or back projection, an

error image is produced by comparing the estimated image with the actual measured data. The algorithm's iteration ends by comparing the initial estimated image with the error image. The initial image estimate represents a non-zero distribution.
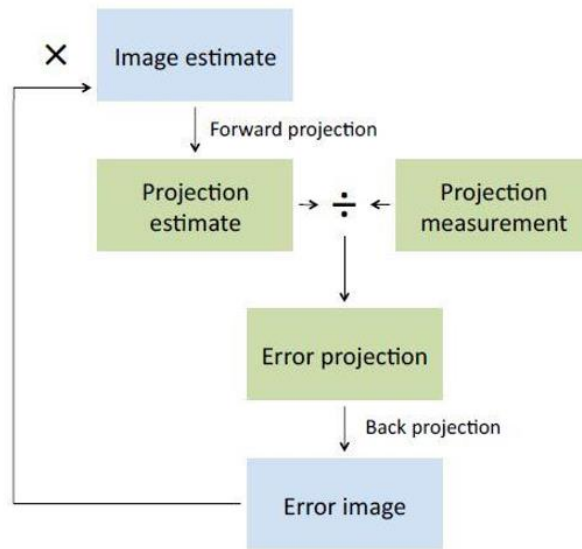


**Figure 4.1:** Schematic overview of an iterative Expectation Maximization (EM) reconstruction algorithm.

So as to obtain realistic results, there are five components that need be defined [15]: (1) a model for the image, (2) a model for the data, (3) a model for the physics of the measurement process, (4) a cost function, and (5) an algorithm to optimize the cost function.

The first (1) is the model for the activity distribution or object density. It represents the image $\lambda_j$ to be iterated throughout the image reconstruction process. The model for the data (2) takes into consideration the randomness of radioactive decay, by reflecting the statistical variation of PET data. For this component, it is considered that positron emission follows a Poisson distribution, so that the data collected is a collection of Poisson random variables. Next, a model for the physics of the measurement process (3) is needed, and this has been introduced as the system matrix. In Chapter 5 we will see how the LORs were calculated and system matrix assembled.

The cost or objective function (4) is the criterion used to determine which image is considered as the best estimate for the object. In the case of statistical algorithms, the cost function is a statistical function. Among these statistical approaches, one can distinguish the Bayesian from the classical methods. The Bayesian criteria for image estimation, such as the Maximum a Posteriori (MAP), assumes the unknown image is random and can be described by a probability density function known *a priori*. Among the classical criteria, there is the Maximum Likelihood, which is of great interest for this work and will be introduced next.

The final component is an algorithm to optimize the cost function (5). When working with Expectation Maximization, the general scheme for the iterative algorithm is displayed in Figure 4.1.

### 4.2.1 Maximum Likelihood Expectation Maximization (ML-EM)

The ML-EM algorithm has become the most common basis for PET data reconstruction approaches. The Maximum Likelihood criterion was first introduced by R. A. Fischer in 1921 [16]. Yet, it was only in 1982 that L.A. Shepp and Y. Vardi [17] introduced a new approach for emission tomography via combining the work of Fischer with a more recent work relating to EM [18]. Their approach took into consideration the Poisson distribution of emission tomography's noise, thus taking into consideration

the characteristic lack of data in PET imaging. The ML-EM algorithm assumes that the quantity to estimate has unknown distribution [19]. Its equation is as follows:

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_i a_{ij}} \sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n} \quad (4.2)$$

Where $\lambda^n$ and $\lambda^{n+1}$ represent the current image estimate and the image estimate that will result in the end of the $n^{th}$ iteration, respectively. $a$ is the system matrix, and $y$ the PET data. The indexes $i$ and $j$ represent the considered LOR and pixel, respectively. In Chapter 5, a practical example will be made to better illustrate how the algorithm works.

The preference of ML estimators, as the one used in the ML-EM algorithm, over other estimators, is based on two reasons related to the concept of bias and variance. ML estimators are asymptotically unbiased because, as the number of observations becomes large, the estimates become unbiased, that is, $E[p(\lambda)] \to \lambda$. ML estimators are also asymptotically efficient because, for a large number of observations, they yield minimum variance, making the ML the estimator least susceptible to noise [20]. Even so, image reconstruction methods based on the ML estimation criterion, like the ML-EM, tend to yield noisy images. This happens since it is in the nature of this estimator to produce images consistent with the data. Since in emission tomography the data obtained is noisy, a good ML image estimate will also be a noisy image.

ᐧThe most common approach when dealing with the ML-EM algorithm is to allow a certain degree of bias in the reconstructed image. This is performed by finding the iteration at which the algorithm converges and, prematurely and intentionally, stop the ML-EM algorithm before it actually reaches the ML solution. Other approaches pass by explicitly implementing spatial smoothing in the images, by using filtering or Bayesian methods, which will be discussed later.

### 4.2.2 OS-EM and GPU Implementations

The main issue when designing an iterative approach to PET data reconstruction is the amount of data that must be dealt with. Alongside, time and computational power constraints lead to the development of algorithm optimizations. These range from developing different algorithms or by using the available software and hardware more efficiently.

A good example for an alternative to the ML-EM algorithm is the widely used OS-EM [21]. The Ordered Subset approach differs from its predecessor by dividing the data in non-overlapping groups. After, standard ML-EM is applied on each subset. Each iteration represents one go at each group, and the image estimate passes on from iteration to iteration. This algorithm is suitable when the data acquired is large enough. If it is not, or if too many subsets are considered, the noise will increase on the resulting reconstructed image.

Regarding using the available hardware more efficiently, efforts are being put to exploit GPU capabilities. GPUs are extremely useful when parallelizing scripted code. They easily surpass CPUs ability to deal with floating points operations [22]. This being, it is useful to parallelize certain parts of the image reconstruction code, as it greatly improves time performance. The main drawback is the code implementation. Lower level programming languages must be used, and the implementation is made to fit a specific GPU card. This means an implementation has reduced re-usage.

## 4.3  Filtered Back Projection (FBP)

Filtered Back Projection (FBP) is an analytic image reconstruction algorithm. The back projection step consists in tracing all the LORs in the spatial domain. Mathematically, FBP is the inverse radon

transform of the sinogram, which was introduced previously. By knowing the system's geometry and having the data arranged in a sinogram manner, the inverse radon transform can be applied to produce the back projection image. However, the back projected image carries a great amount of noise. The main source for this noise comes from the loss of high frequencies when converting from the Fourier to the cartesian space [23]. The common procedure to contradict this effect is to convolute a filter with it, hence producing a Filtered Back Projection image.

Image reconstruction using FBP yields great results, if taken into consideration it is computationally inexpensive. However, even with filtering, this method continues to yield great amounts of noise, especially when dealing with the small statistics characteristic of PET technique. Yet, this is taken as the standard image reconstruction method and, therefore, will serve as a comparison to the image reconstruction algorithm developed in this work's scope.

# 5 Data Correction and Regularization Methods

Images in Emission Tomography techniques are associated with high noise. The small data, photons interaction with matter, and hardware limitations, are the main factors for noise presence. Different noise reducing algorithms or quality enhancement approaches can be used to correct the data acquired, and preserve or enhance image quality. Exploiting different hardware's capabilities can also contribute to noise reduction. Noise presence and overall image quality can be represented via signal-to-noise ratio (SNR) or through analysis of spatial resolution, among others.

## 5.1 Data Corrections

Data affecting events can be compensated through statistical approaches. Among these, some relate to the photons, such as: attenuation and scatter, TOF, and false coincidences. Hardware and reconstruction algorithms also affect noise presence by either reducing or incrementing it [7], [24].

### 5.1.1 Attenuation Correction

Taking into consideration photon attenuation enables correction of enclosed tissue regions which are falsely reconstructed with low activity. Photon intensity attenuation is given by:

$$I(x) = \frac{I_0}{e^{\mu x}} \qquad (5.1)$$

Where $I(x)$ is the intensity of the beam after traveling a distance $x$ in a given material with attenuation coefficient $\mu$. $I_0$ is the initial beam intensity. By taking into account this effect, and using an attenuation map, one can say that if a given LOR passes through more attenuating tissue, its corresponding pair of detectors should have, statistically, detected more coincidences. This correction is often used in human PET scans, where the detector's size is considerable and there is significant tissue attenuation. The attenuation map is often a transmission image of the object, and represents an attenuation coefficient distribution map [8].
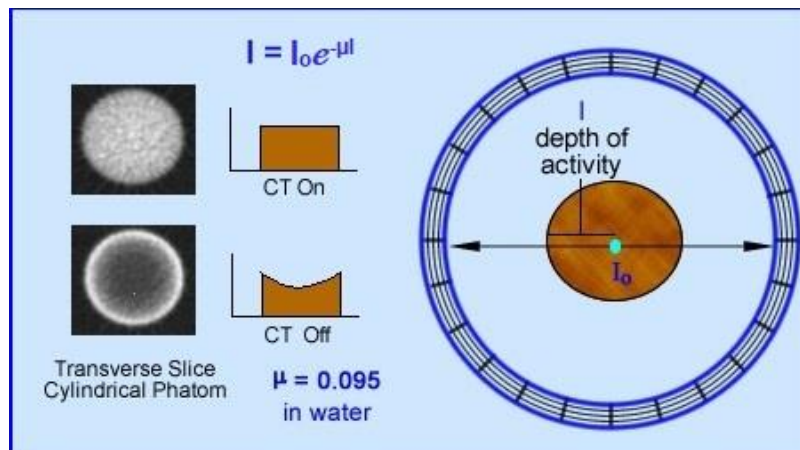


**Figure 5.1**: Representation of the effect caused by photon attenuation, with and without attenuation correction *[25]*.

### 5.1.2 Time of Flight (ToF)

Time of flight is the precise measure of the time interval between the detection of both coincidence photons. By measuring this variable, we can obtain a better statistic distribution for the location of an event, as it indicates to which detector the event has occurred closer to (see Figure 5.2). This correction method requires hardware with high time resolution capabilities, with modern detectors having timing resolutions between 580-700 ps, sometimes as low as 300 ps [26].
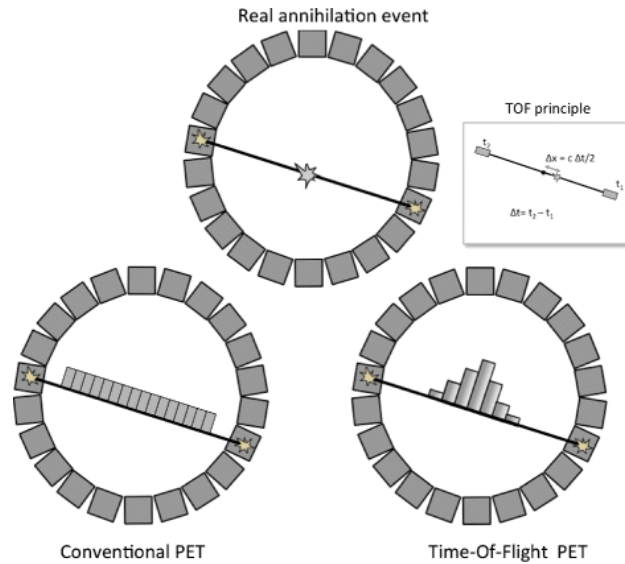
**Figure 5.2:** Compared to conventional PET, the estimated ToF difference ($\Delta t$) between the arrival times of photons on both detectors in TOF-PET allows localization (with a certain probability) of the point of annihilation on the LOR. In TOF-PET, the distance to the origin of scanner ($\Delta x$) is proportional to the TOF difference via the relation $\Delta x = \frac{c\Delta t}{2}$, where $c$ is the speed of light, $t_1$ the arrival time on the first detector, and $t_2$ is the arrival time on the second detector *[26]*.

### 5.1.3 False Coincidences – Scatter and Random Coincidences

Scatter events are responsible for activating detectors that do not include the initial photons directions. This leads to a mismatched LOR being considered. A common approach is to define a narrower energy window for the detected photons. As scatter results in the reduction of the photon's energy, it is possible to reduce the contamination caused by this effect. In this case, the detector's crystal should present good energy resolution capabilities.

Much like scatter events detection, random coincidences also produce mismatched LORs. However, limiting the energy window is not effective, as these coincidences can be produced by two unscattered photons. As is the case of TOF time resolution requirements, here goes the same. For a coincidence detection to be triggered, two detectors have to be paired. If the time interval between both triggers is too long, both photons do not belong to the same annihilation event. As can be seen in Figure 5.3, if the time interval between two detected photons is longer than the maximum time it would take the photon to cross the full FOV, then the coincidence is either random or results from a scatter event. Either way, the coincidence must be dismissed.
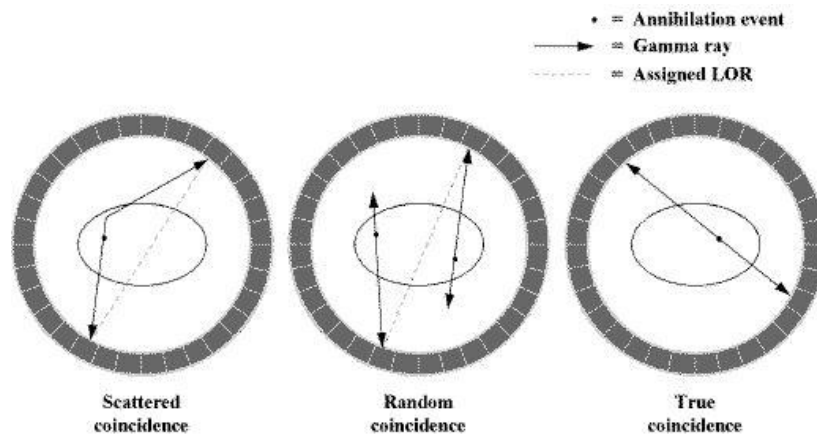


**Figure 5.3:** Type of coincidences in PET *[27]*.

The hardware also takes a role in contributing to noise reduction. As was mention previously, detectors with high temporal and energy resolutions are useful when aiding in noise reduction. However, detectors have inherent problems related to their actuation and uniform behavior. The solutions for these problems appear, among others, as dead-time correction, and detector normalization [28].

## 5.1.4 System's Geometry and Actuation

The dead-time is the time interval after a photon arrival in which a detector remains idle and cannot process other events. It results in smaller data counts and in the saturation of detectors. Knowing the dead-time inherent to a system allows the correction of saturated detectors, as these loose counts that would render them with much more statistic than those of unsaturated ones.

Detector normalization allows correction of LORs' sensitivities. Sensitivity is affected due to the geometry of the system and hardware constraints [29]. In the first case, the angle to which a LOR intersects each of the detectors face. relative to the mean incident angle (see Chapter 5.1.5), strongly affects the sensitivity. The wider the detector's face is, the more this effect can affect sensitivity. As for hardware constraints, the efficiency throughout all detectors is not always the same. In a block of detectors, there can exist a heterogeneous distribution of gains, which leads to sensitivity variability. In both cases, a full scan where all possible LORs are activated, can be performed. This method allows retrieving information on LOR sensitivity, and acquire the normalization coefficients for each LOR.

## 5.1.5 Depth of Interaction (DOI)

Lastly, one can have detectors able to measure the Depth of Interaction (DOI) [30]. The DOI is the point in the detector's crystal where the photon interacted. This metric is possible through the double readout present in more advanced detectors, which allows the measurement of the asymmetry in the collected light at both ends of the detector's readouts. As when tracing a LOR one would previously draw it from the middle point of both detectors face, the DOI error correction allows retracing LORs affected by this parallax error by adding information on photon detection. This way, it is possible to retrace the LOR in a more correct manner. Adding DOI correction into a system allows the improvement of radial resolution, as shown in Figure 5.4.
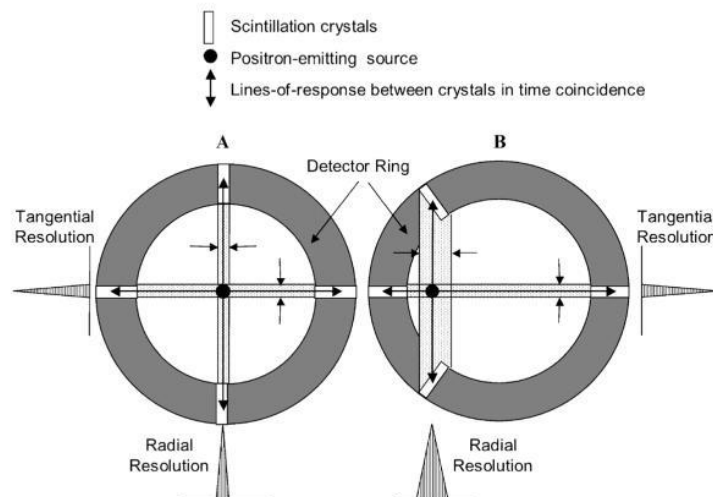


**Figure 5.4:** A) For a point source near the center of the FOV, photons enter crystals in the detector array through their very small front faces and the difference between the LORs and the true photon flight paths is small, i.e. results in "good" radial resolution. B) For off-axis sources, photons can enter the crystals through their front faces and anywhere along their sides, so radial resolution is "poor". Note that tangential resolution is not dependent on the DOI effect and is essentially constant across the FOV *[31]*.

## 5.2 Data Regularization Methods

Preemptive noise reduction at machine and acquisition levels is very important. Still, although image reconstruction methods are designed to better fit the characteristics of the data and the user's necessities, the statistics of PET data is often scarce, and image reconstruction will always carry noise. To minimize noise presence, quality enhancement methods are implemented throughout the image reconstruction step. These methods consist mostly on statistical approaches that aim to reduce noise by taking into consideration the local behavior on reconstructed images. What follows is an overview on how these methods actuate and the types that are of most interest to this work.

**Figure 5.5:** Noise presence in sources combined with poisson noise to demonstrate the noisy nature of PET imaging. The images shown correspond to a) 10 counts per pixel; b) 100 counts per pixel; c) 1000 counts per pixel; d) 10000 counts per pixel.

### 5.2.1 Inter-Iteration Gaussian Filtering

Image filtering can be performed in the spatial or frequency domains. Filters are used to emphasize or remove certain image features, by performing smoothing, sharpening, and edge enhancement. Filters that operate in the frequency domain are of most use when dealing with frequency related problems, for example, removing the Mains Hum (power grid current frequency) from any data acquired with a system sensible to it.

Images can also be filtered in the spatial domain. This filtering approach consists in considering a neighborhood around a given pixel, so that its neighborhood has some weight on the new value for it. For each method, the dispersion and size of the neighborhood, as well as the weight given to it, must be decided. In this work, Gaussian filtering was considered.

Spatial Gaussian Filtering consists in using a Gaussian distribution as a "point-spread" function, which is achieved through convoluting it with the image. Since a pixelized image is a discretization of an object into pixels, a discrete approximation to the Gaussian function must be made. Since the Gaussian distribution is non-zero everywhere, its discretization would create an infinitely large convolution kernel. However, in practice, it is approximately zero at more than three standard deviations and this is often used as a kernel cut-off point [32].

### 5.2.2 Median Root Prior (MRP)

Other than using filters, image quality enhancement can be made based on *a priori* knowledge about the nature of the image. Such approaches guide the image reconstruction process into solutions considered more favorable. Among these, Bayesian inference is of great use. The logic behind Bayesian statistics states that the knowledge on prior events can be used to better predict future ones. In the case of image reconstruction, priors can be seen as the knowledge on the nature of the image. In the scope of this work, we will introduce the Median Root Prior.

The median root prior (MRP) [33] is based on the general assumption that an ideal PET image consists of constant neighborhoods with monotonous transitions between them. These are also the characteristics of the root signal of a median filter. A root signal remains unaltered when its corresponding filter is

applied to it [34]. As the median filter presents good results in noise reduction and edge preservation, it is expected that an application of a prior based in it should also yield similar results. When implementing the prior into the ML-EM expression, the algorithm expression is as follows:

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_i a_{ij} + \boldsymbol{\beta} \frac{(\boldsymbol{\lambda_j^n - M})}{\boldsymbol{M}}} \sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n} \qquad (5.2)$$

Where the new term, in bold, is comprised of a parameter $\beta$, defined by the user, and the median value $M$ of the considered neighborhood of $\lambda_j^n$.

Contrary to other priors, the only parameter needing adjusting when using the MRP is the $\beta$ parameter. This reduces any other parameter definition or optimization. One other intuitive alternative would be using an average, yet this results in blurred images, whereas the repetitive application of the MRP produces an unaltered root image [35].

## 5.3 Image Resolution and Quality Assessment

It is important to stress that image quality is affected not only by the points explored above. The data sample should be large enough, and the number of pixels used to discretize the image is also very important. Using more pixels will lead to noisier images, as the detectors and data are limiting agents to spatial resolution. However, the pixel size must be set accordingly, as the sizes of the structures one is looking to observe must not be smaller than the pixel itself.

Image quality assessment is of great importance when comparing the effectiveness of the implemented reconstruction or enhancement methods. For this, it is important to use standardized methods that ensure normalized quantitative results in order to compare different methods or imaging systems. Some of these quantities are the Signal to Noise Ratio (SNR) and spatial resolution.

### 5.3.1 Image and Spatial Resolution

In imaging, spatial resolution translates the ability to separate points that lie close to each other. This is also known as the minimum resolvable distance. For PET imaging, this quantification allows a comparison between systems, reconstruction methods, and image enhancement techniques. Through several trial scans using phantoms and controlled environments, it becomes possible to determine the spatial resolution of a given system, and use it as a reference in clinical practice.

Image resolution also translates the ability to separate points that lie close to each other. Yet, this refers to the ability of a given peripheral to produce separate points. Whereas spatial resolution depends on the information available and is limited by the scanner characteristics, image resolution depends on the number of pixels or discretization units one can produce. With image resolution comes image quality, which translates the capability in storing and displaying raw information, without needing to compress it.

Altogether, image resolution can be a limiting agent to spatial resolution, since even though the information might be present, the displaying system might not be able to present it properly. Or the contrary, where a displaying unit has good image resolution, but the data has reduced spatial resolution.

# 6 PET Systems

PET imaging systems are designed to fit a wide range of applications related with medicine. Systems can be designed for didactic, pre-clinical, or human scanning purposes. This being, it is expected that these systems present different characteristics that best suit their purpose. They can vary in the geometry, hardware, reconstruction method used, and how the object is imaged, either single-sliced, multi-sliced, or 3D.

## 6.1 Didactic, Pre-Clinical, and Human Scanners

Didactic PET scanners allow field related students and novice practitioners to train with this technique, and better understand how it works. As clinical and pre-clinical systems are out of most education organizations budget, designing cheaper systems allows the opportunity to close the learning gap between theory and practice. Didactic systems disregard having great image quality, hence lowering the number and quality of high cost hardware components, ultimately reducing the system's cost. Rather, the didactic approach passes by enabling observing the actuation of simpler electronics and mechanics, as well as easily implementing and experimenting different image reconstruction and enhancement methods. Eventually, these systems will lead to more interested and qualified professionals to work with pre-clinical and clinical systems.

On the other end, we have clinical PET systems, which are present in medical care facilities. These systems are designed to meet practitioners needs and wants, leading to higher initial and operating costs, which can surmount several hundred thousand dollars. The associated costs cover the expensive hardware and all the R&D process behind the development of a given machine.

Finally, there are pre-clinical scanners. These scanners are often designed for small animal's research due to their high importance when testing new pharmaceuticals meant for humans. One would think the smaller size of pre-clinical scanners, when comparing with clinical systems, would mean a price reduction. However, small animal's scanners use smaller detectors and have a more compact assembly. This is to keep spatial resolution the size of the smaller biological structures of interest. The smaller detectors, with the same actuation characteristics of those found in clinical scanners, are more expensive to produce, as well as the electronics associated with them. As a result, the price of the system increases.

In this work, the image reconstruction algorithm was implemented for two 2D systems, both based on the same geometrical design. The system is called easyPET and it was first thought of as a didactic PET system. With the advancing of the project, the idea of designing a small animal's scan for pre-clinical applications rose. Hence, the easyPET for pre-clinical application appeared.

## 6.2 Scanners Geometries and Dimensional Acquisition

Scanners are designed in a wide variety of ways to better fit their purposes. Usually, the machines are also built for the user to have some freedom when planning a scan. A scan is limited by the scanner's geometry, the resulting image, and, as was mentioned previously, is also affected by the technique's inherent properties and hardware constraints.

The geometry of a scanner is the result of an extensive research on its goals and limitations. Usually, scanners are designed with circular geometry, on which the modules of detector arrays are fixed at a given distance from the center of the field of view (FOV), forming a ring. Considering that the ring is defined in the xy plane, there is a set of consecutive rings along the z axis. This type of geometry is recurrent because of small animals and human's geometries, which can easily be aligned with the z axis.

Obviously, one could design PET scanners meant for specific body structures, as is the case of Positro Emission Mammography (PEM) scanners. Yet for the sake of reusability, these scanners are the most versatile, and work well with fairly centered to the FOV and round structures like the heart, brain and neck, among other organs.
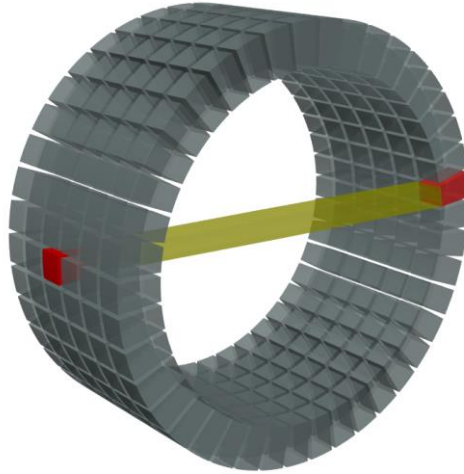


**Figure 6.1:** Representation of a 3D ring assembly of PET detectors *[36].*

In terms of the data to be reconstructed, systems can output data that corresponds to the object seen in 2D/planar or 3D/volumetric images. Planar images are reconstructed only with LORs detected in the same xy plane of a considered slice. Each reconstructed slice has a thickness corresponding to the detectors side. Reconstructing an object with multiple 2D images results in a 2D multislice image. Finally, 3D image reconstruction implies taking into account LORs that cross more than one xy plane. Volumetric images are composed of voxels, which are volumetric pixels. The imaged space can be discretized by voxels with a thickness the same as the detectors, resulting in a 3D multislice image, or with cubic voxels with a face area smaller than the detectors, resulting in a higher resolution 3D image.

## 6.3  State of the Art

Next, a brief state-of-the-art will be made on four pre-clinical PET scanners. These systems are presented here as examples for the current hardware, image reconstruction methods, and achieved results, even though there are several other systems worth mentioning.

### 6.3.1  DigiPET – MEDISIP/INFINITY

The DigiPET system is a scaled down system, using only four modules of 32x32 mm$^2$ detectors placed in a square arrangement (see Figure 6.2). Opposing detectors are 34.5 cm apart, and the system produces a FOV of 32x32x32 mm$^3$. This system uses LYSO crystals and digital silicon photomultipliers (dSiPM). It has several image reconstruction algorithms implemented, being its predefinition the 3D-OSEM with ray tracing using the Siddon's algorithm [37].
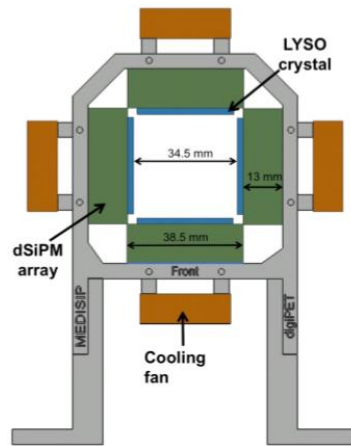
**Figure 6.2:** Schematics of the DigiPET system prototype *[37]*.

## 6.3.2  β-cube – Molecubes/IBiTech/MEDISIP/INFINITY

This system comes from a consortium between the entities that produced DigiPET and a spin-off from Ghent University. It has 9 detector modules arranged as a ring with 76 mm diameter, and an axial FOV of 13.3 mm (see Figure 6.3). Each module contains five 25.4x25.4*8 mm$^3$ thick monolithic LYSO scintillators coupled to analogue SiPMs. This system has a GPU-based Tube of Response 3D-OSEM implementation. The main characteristic about $\beta$-cube is that it is designed to be coupled with a CT scanner, *X*-cube, turning it into a bimodality scanner [38].
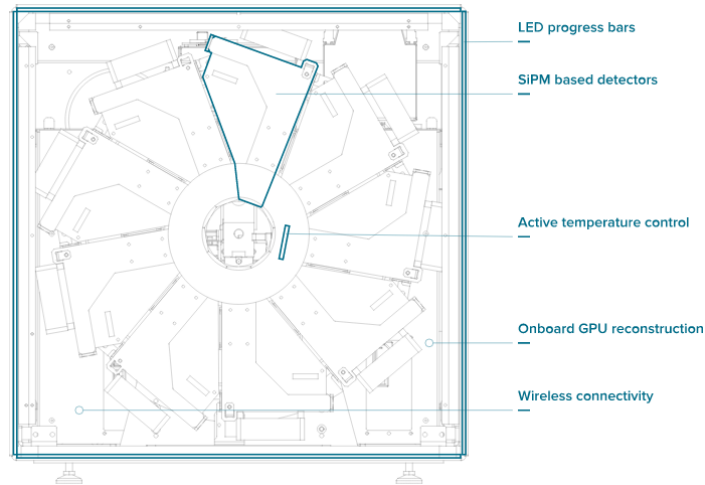


**Figure 6.3:** Schematic axial cut of the $\beta$-cube *[38]*.

## 6.3.3  NanoScan PET/MRI – MEDISO

The NanoScan PET/MRI is, as the name states, a bimodality scanner. For the PET section, it has an axial FOV of 94 mm. Each detector module can be placed in coincidence mode with 1, 3, or 5 other modules, allowing for a transaxial FOV of 45, 94, or 120 mm, respectively. The PET detector consists of 12 modules, each a 39x81 mm array of 1.12x1.12x13 mm of LYSO crystals. The system uses position sensitive photomultipliers (PS-PMT). Since this PET scanner is coupled to an MR scanner, it is reinforced with an internal radiofrequency shield. This system also uses a 3D-OSEM implementation [39].
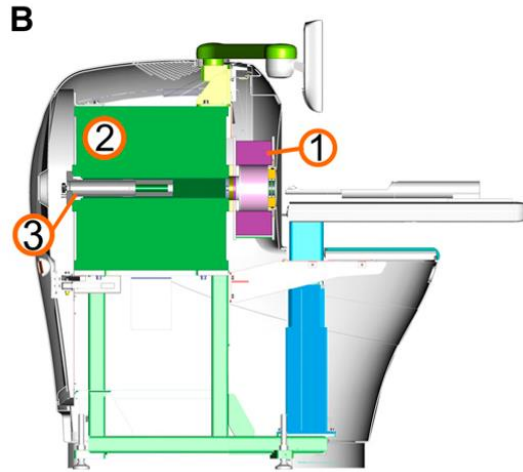
**Figure 6.4:** nanoScan small-animal PET/MR scanner schematics. Labeled components are: (1) PET ring; (2) magnet; (3) radiofrequency coil. The increased size concerns to the difficulty in combining these two techniques *[39]*.

### 6.3.4 Albira PET/SPECT/CT – MILabs

Finally, the Albira system is composed by eight removable modules, arranged in an octagon. It has a transaxial FOV of 80 mm and axial of 40 mm (see Figure 6.5). It uses non-pixelated LYSO scintillators with 10 mm thickness. Connected to these are multi-anode photomultipliers (MA-PMT), which allow for DOI measurement. As the aforementioned systems, the Albira also has a 3D-OSEM implementation [40].
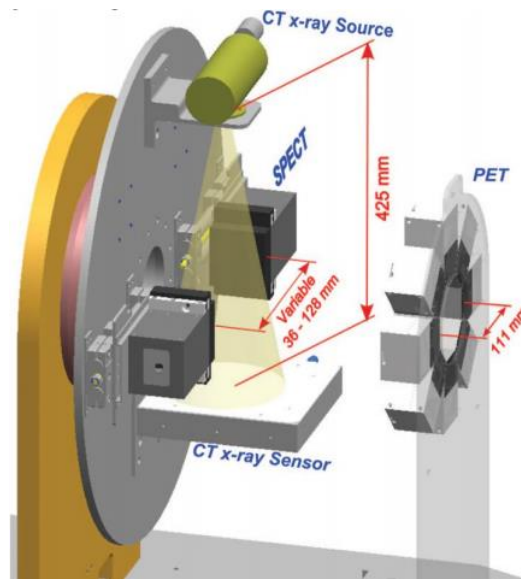


**Figure 6.5:** Schematic view of the entire Albira PET/SPECT/CT system *[40]*.

**Table 6.1:** Summary of the systems presented above. *Algorithm used for spatial resolution measurement following NEMA rules. **Native algorithm developed for the system.

| SYSTEM | SPATIAL RESOLUTION (MM) | IMAGE RECONSTRUCTION USED |
|---|---|---|
| DIGIPET | 0.7 | MLEM*/3D-OSEM** |
| NANOSCAN PET/MRI | 1.5 | FBP*/3D-OSEM** |
| $\beta$-CUBE | 0.780 | FBP*/3D-OSEM** |
| ALBIRA PET/SPECT/CT | 1.65 | FBP*/3D-OSEM** |

# 7 easyPET System

## 7.1 Project and Work Group

As was previously mentioned, the system used in this work was easyPET. Currently, this system has two variations. One being for didactic purposes, and the other with pre-clinical applications for small animals. The easyPET was initially designed as the first afore mentioned variation. Its reduced number of detectors alongside the innovative geometry, resulted in an affordable PET scanner for educational purposes [41], [42].

The didactic easyPET scanner was first idealized at University of Aveiro, and the work was developed alongside high school students and Science centers, so as to spread the knowledge involved in PET technique. The patent for this system is now part of Caen [43], which is now selling the system. With the potential seen in this innovative approach, efforts were put to develop the scanner towards having pre-clinical applications. Also, a spin-off, RI.TE [44], was created around the project. Consequently, a work group was formed between i3N Lab [45], RI.TE, University of Aveiro, University of Coimbra, University of Insubria, and University of Lisbon.

The pre-clinical scanner project is still ongoing. A 2D scanner is already assembled and working, and its corresponding image reconstruction methods were developed in the context of this work and are shown in this thesis. Currently, the 3D mode is being developed, already with a working prototype. In a later phase, tests with small animals will be performed. Note that throughout all these steps, image analysis is always needed.

## 7.2 Concept Geometry

The geometry designed for the initial patent can be seen in Figure 7.1. To produce a slice, the system only relies on two facing detector arrays (3 and 4), at a distance d from each other, and two rotation axes ($\alpha$ and $\theta$). During the revolution (Figure 7.1 - Left), the center of the FOV (1) defines the bottom/$\alpha$ angle axis, and both detectors revolve around it, performing a full 360º revolution in a certain number of $\alpha$ steps. At each $\alpha$ step, a fan-like motion takes place (Figure 7.1 - Right). The detector (4) revolves around the detector (3) with a user defined amplitude and in a certain number of $\theta$ steps. The detector (3) rotates around its axis so that it is always facing the detector (4). The top/$\theta$ angle axis produces a fan-like movement.
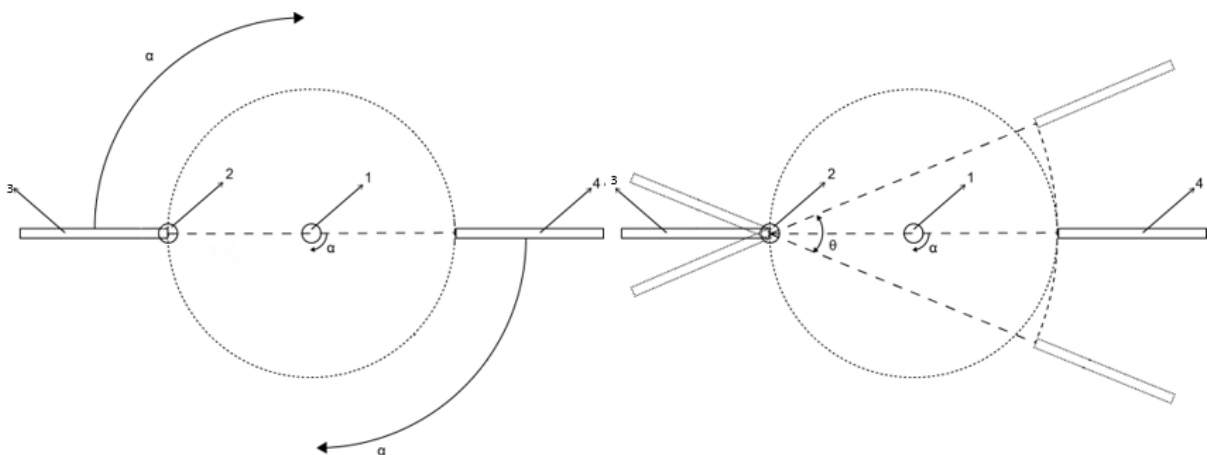


**Figure 7.1:** Schematic for the easyPET geometry, retrieved from the patent file. (1) corresponds to the bottom/$\alpha$ angle axis of rotation; (2) corresponds to the top/$\theta$ angle axis of rotation; (3) and (4) correspond to the two detector arrays. Note that the top/$\theta$ angle axis of rotation is fixed in (3), more specifically the detector's face. **Left:** schematic for a bottom/$\alpha$ angle revolution.

**Right:** schematic for the fan-like movement, defined by the top/$\theta$ angle, performed at each bottom/$\alpha$ angle step. Adapted from the easyPET system's patent.

As one can observe, the fact that both detectors are always facing each other, renders this geometry the characteristic of not having DOI related problems. The reduced number of detectors also confers a diminished probability of detecting random coincidences and scatter events. On the other hand, this will eventually result in reduced statistics.

## 7.3 Hardware and Actuation

The system comprises a U-shaped circuit board, with each detector facing each other, exemplified in Figure 7.2. The SiPM detectors vary between 2x2x30 mm$^3$ and 1.5x1.5x30 mm$^3$, and they are 57.7 mm and 60 mm apart, respectively.

The system has two motors that control the rotation of both axes. The motor that controls the $\theta$ axis is the top motor, and the one that controls the $\alpha$ axis is the bottom motor. Both motors and event detection were initially controlled via an Arduino. However, due to several limitations in using it, a new microcontroller was implemented. This has allowed a better control on the motors actuation and greatly reduces dead-time.

Via a user interface, the user is able to control the acquisition parameters, such as time spent per step and step size, the number of $\alpha$ revolutions, the fan-movement amplitude, as well as event detection related parameters like energy cut-offs.



**Figure 7.2:** Image used for commercial purposes by Caen, portraying the U-shape PCB with two detectors, one at each U-tip [37].

## 7.4 3-Dimensional Acquisition Mode

The description above was made regarding the system in 2D mode. The multislice and 3D modes model are described very similarly. Instead of a single pair of detectors, this model includes a pair of column arrays of detectors oriented in the z axis direction. Aside from allowing 3D or multisliced images, the main differences with multiple paired detectors are the increased number of events detected, which can lead to actuation constraints, and the slightly increased probability for random coincidences. Even though, the image reconstruction implementation is very similar in all three modes. However, the 3D implementation still requires taking into consideration the spatial domain with one more degree of freedom, which greatly increases the algorithm's implementation complexity. The multislice and 3D modes are designed to be used in the pre-clinical easyPET scanner, depicted in Figure 7.3.
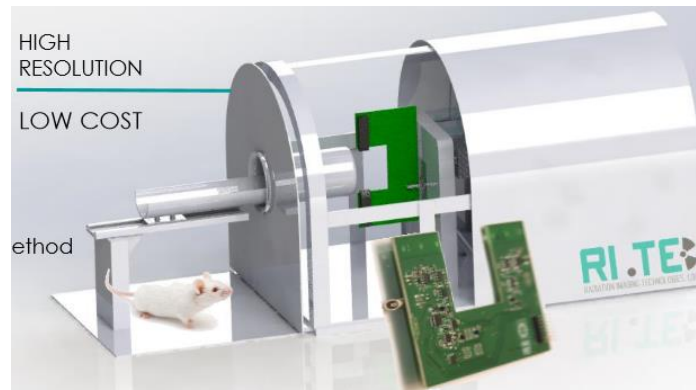
**Figure 7.3:** Design for the pre-clinical easyPET scanner *[44]*.

# 8 Motivation

This thesis is mainly centered in the work performed implementing the ML-EM iterative algorithm for the 2D easyPET system, as well as the optimizations made. It also aims to display the results that this implementation has produced and the benefits of using data regularization methods. Moreover, this thesis will hopefully showcase the easyPET system and its potential, in addition to some of the knowledge gathered when working inside the easyPET project and close to the PET technique. It was also of great importance to produce a tool, in the form of a user interface, so as to aid the didactic easyPET system in better performing as an educational tool.

# 9 Materials and Methods

In this chapter is presented the thought process behind taking into consideration the easPET system's geometry and data organization, and implementing the ML-EM algorithm with Gaussian filtering and Median Root Prior for image quality enhancement, using the licensed Matlab software. The main optimizations to the program code will also be presented and discussed. Following this are the procedures used for image analysis.

## 9.1 Matlab software

The Matlab software is widely used by the scientific and academic communities. Its popularity comes from the wide variety of applications, active community, and support [46]. Being in constant development and improvement, it has an extensive library, covering Mathematics and Statistics, Physics Modeling, Signal and Image Processing, among several other topics. Each topic is divided into toolboxes comprising several functions that relate to a given subject (e.g. the Signal Processing Toolbox allows filter design, amongst others). The user may also design its own functions, or recur to ones submitted by the community.

At a user level, one encounters a C-like language, yet with lower complexity. In many cases, this only demands the user to arrange the available functions according to the problem he is trying to solve. Despite not being a low-level programming language, Matlab still requires an advanced knowledge on the logic of implementing a problem in a programming language. One of the most attractive Matlab particularities, are the multiple user-friendly interfaces with which the user can interact to utilize the available functions. On top of that, Matlab offers a Guided User Interface (GUI) creator, called GUIDE. This application is very important when one of the goals of a project is for the solutions achieved to be available in a user-friendly fashion. In the case of the work presented here, a user interface was created to better lead people when dealing with image reconstruction for the didactic easyPET system.

For this work, most of the scripted Matlab code does not comprise native functions, due to the singularity of this problem, and the liberty it conveys when optimizing the solution implementation. The few recurring functions are used, solely because they do not present optimization constrictions, and can be easily dealt with in the future, so that effort was not put into removing them. Also, in the context of the ongoing project for the didactic and pre-clinical scanners, it is of great interest to, in the future, implement the work presented in this thesis using a freeware coding language, such like C or C++. Any of these languages is widely spread, more computationally efficient, and the system's software availability would not be limited by any license fee. This being, it remained important to leave the use of native Matlab functions at a minimum, as it would pose as additional work when implementing them in the new language.

## 9.2 easyPET system acquisition

In order to acquire or simulate PET data, the user has to define certain parameters that control energy cut offs, time of acquisition/simulation, amongst several others. However, contrary to most scanners, the easyPET is comprised of moving detectors (see Chapter 7) instead of the common static detector placement. This demands the user to input additional parameters that will define the behavior of the detectors across the system's geometry. Here on after, the $\theta$ angle will be referred as top angle, and the $\alpha$ angle as bot angle. This arrangement arises from the fact that the $\theta$/top angle is controlled by the top motor, and the $\alpha$/bottom angle is controlled by the bottom motor of the easyPET system.

Considering the detectors' movements, the additional parameters needed for an easyPET scan are: the bot and top angles step, the time spent by top angle step, the top angle range, and the number of revolutions the system is to perform. The bot angle range is not mentioned, as it is considered that the system will only perform full revolutions, and at least one, so that the bot angle always ranges from 0º to 360º, and all revolutions produce the same LORs.

Since the range of the bot angle is set to be 0º to 360º, defining the bot angle step determines how many bot angle values there will be in a single system's revolution.

$$Number\ of\ BotAngleValues = \frac{360}{BotAngleStep} \quad\quad (\ 9.1\ )$$

At each bot angle value, the system performs a fan like motion. The minimum and maximum top angle values limit this motion's range. Usually, these values are set to be symmetric, but they do not need to be. Much like the bot angle, the top angle step determines how many steps there will be per fan like motion. In this case, contrary to a full revolution, it is necessary to add 1 to correct the missing step.

$$Number\ of\ TopAngleValues = \frac{TopAngleMax - TopAngleMin}{TopAngleStep} + 1 \quad\quad (\ 9.2\ )$$

This expression gives us the number of top angle values per bot angle value. Since a LOR is described by a unique combination of a bot and top angles values, given the expressions 9.1 and 9.2, the total number of unique LORs liable to be produced in a given scan is:

$$Number\ of\ Bins = Number\ of\ BotValues * Number\ of\ TopValues \quad\quad (\ 9.3\ )$$

The time spent per scan is easily obtained. Through the defined time per top angle value and the number of revolutions performed, we have:

$$Time\ Per\ Scan = Number\ of\ Bins * Time\ per\ Step * Number\ of\ Revolutions \quad\quad (\ 9.4\ )$$

## 9.3 easyPET Data

### 9.3.1 List Mode

As was introduced before, most PET scanners output their data in the form of a list mode file. The easyPET raw data is outputted as a .txt file with a number of entries corresponding to the number of photons detected. Each entry has two labels, one corresponds two a time stamp, the other to the detector triggered. The list mode can be processed afterwards, resulting in a clearer data file.

### 9.3.2 Processing the List Mode

#### 9.3.2.1 Pairing Entries

After acquiring the PET data as a list mode file, one must process it into a file that translates the spatial coordinates of the active detector pairs and their count values. As the easyPET only has two detectors, it becomes easier to do this step by just looking at the time stamp of two consecutive list mode entries. If both entries belong to a different detector and are found to be within a reasonable time interval, as was discussed previously (Chapter 3.1.1), both detections result from the same annihilation event and the correspondent bin increments in one count.

#### 9.3.2.2 Spatial Coordinates

Contrary to scanners with static detectors, where the location of detector pairs is independent of time, the easyPET has its detectors constantly performing a step-by-step motion defined by the angle steps

and time per top angle step, always according to its geometry. By knowing the time per top angle step defined for a given scan, one can deduce the detector location at each time stamp. Hence, we can deduce the bot and top angles of a given LOR, at a specific elapsed scan time.

### 9.3.2.3 Post-Processed Table

For the image reconstruction to begin, it is necessary to have a data file that results from a processed list mode, and, much like the sinogram, this table represents a histogram of the data, but with explicit representation of bin coordinates. This processed data file results from taking into consideration what was discussed in Chapter 3.1.1. There is no standardize organization for the data at this stage. However, depending on the image reconstruction method or available freeware (e.g. STIR [47]), one can require a predefined data organization, such as a sinogram.

In the case of the 2D implementation for the easyPET system, the used data file was organized having three columns. The first corresponding to the bot angle values, the second to the top angle values, and the third to the accumulated counts corresponding to the bin described by the bot and top angles values. This file has an entry for each possible detector pair, despite not all having associated counts.

### 9.3.2.4 Sinogram

As was mentioned previously, the sinogram is the most traditional way of displaying processed PET data, since it relates to the Radon transform. By simply applying the Inverse Radon transform on the sinogram, one can obtain the back projection of tracer distribution. To be used with Filtered Back Projection (FBP), a sinogram can be organized in several ways, provided that the Inverse Radon Transform is in concordance with it. Traditionally, a sinogram is built as follows.

A Line of Response (LOR) in a 2D PET scanner may be specified by a set of two coordinates $(s, \phi)$, where $s$ is the transaxial distance from the LOR to the scanner axis and $\phi$ is the azimuthal angle of the LOR (see Figure 9.1). Then, each bin of the sinogram matrix will have the count value correspondent to the detector pair described by the specific set of coordinates.
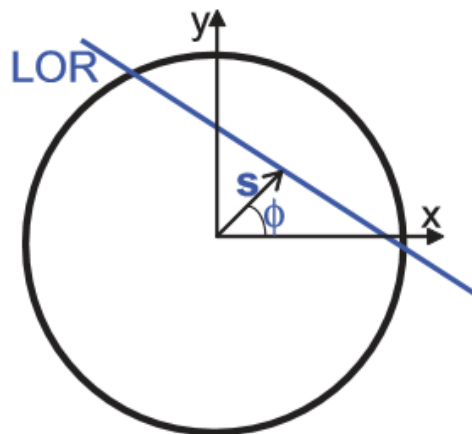


**Figure 9.1:** View of a LOR projected into the transaxial plane, where the LOR is described by the coordinate pair $(s, \phi)$. Adapted from *[48]*.

## 9.4 LOR and System Matrix

### 9.4.1 LOR

The Line of Response is the basic structure for PET image reconstruction. It allows converting the data obtained in a PET scan into the spatial distribution of radiotracer. Due to the nature of computerized images, which discretize space, LORs are too represented as a discretized portion of space that translates

the probability distribution of a given annihilation. For this work, LORs were computed using a ray-driven approach.

### 9.4.1.1 Ray-Driven Approach

Ray-driven approaches rely on taking a detector pair, and connect them with a line, known as LOR. As this line is computed in a discretized image space, it is necessary to calculate the LOR intersection with a x-y grid that represents the pixels in the image (see Figure 9.2) [49].
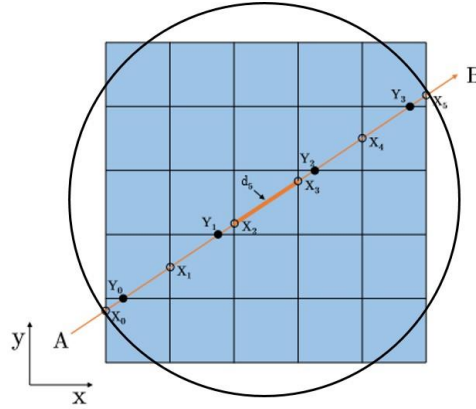


**Figure 9.2:** Representation of the ray-driven method. For simplicity, each detector's face's middle point is considered in contact with the FOV.

For this, one must first have two points in space, which correspond to the mid points of each detector's face. Next, we can calculate the slope, $m$, of the line, and its intersection, $b$, with the y axis. As so, we obtain the reduced equation of the LOR.

$$y = mx + b \qquad (9.5)$$

To calculate the line segment that crosses each pixel, we first calculate the x and y values correspondent to each LOR intersection with the x-y grid defined by the pixels. After, we can finally calculate the length of each pixel intersection with the LOR. Knowing that a LOR as a maximum probability of one, the sum of all intersected pixels must also equal one. Each pixel is assigned with a new value corresponding to:

$$New\ pixel\ value = \frac{Pixel\ line\ segment\ length}{LOR\ length} \qquad (9.6)$$

The described ray-driven approach was implemented as a Matlab function. This function takes as input the number of pixels the imaged space is to have, and the top and bot angles values of the detector pair that produces the desired LOR. This function is called during the construction of the system matrix, which will be described next.

### 9.4.2 System Matrix

The system matrix comprises all the LORs that can be produced in a scanner. Once the ray-driven method was implemented as a Matlab function, this was easily achieved by running this function multiple times, as many as needed to run through all possible combinations of top and bottom angles. Each time the ray-driven function is run, it outputs a LOR corresponding to a pair of top and bottom angles. The output is then assigned to a larger matrix, organized according to the sequence of top/bottom angles pair in the data file.

A Matlab script was designed as a function to build the system matrix. This script takes as input the range of the top angle values, in the form of a minimum and maximum values, and the top angle step. It then creates an array with all possible top angle values, and runs the ray-driven method for each possible LOR. The LORs in the system matrix are organized sequentially, meaning that the first LOR corresponds to the first bin in the data file, the second LOR to the second bin, and so on and so forth. However, system matrixes are computationally heavy. For example, a system matrix created to reconstruct 100x100 pixels images, with 32800 possible LORs (10000x32800 entries), and in sparse mode, takes up to 64 MB of RAM when running the reconstruction algorithm. This characteristic has led to the appearance of several approaches, such as using sparse matrixes, performing the ray-driven method "on-the-go", or using geometry symmetries. What follows is the usage of sparse matrixes and the exploitation of easyPET's geometry symmetries, in order to reduce the computational heaviness of the system matrix.

### 9.4.2.1 Sparse Matrixes

A matrix is considered sparse when most of its elements are zero. A computed LOR in the imaging space corresponds to a very small amount of non-zero pixels. Therefore, the system matrix is a collection of smaller sparse arrays. Matlab enables reducing the amount of allocated memory in sparse matrixes by reducing the memory allocated to zeros. The more sparser and greater a matrix is, the more beneficial it becomes to use this approach. This is performed by using the native Matlab function sparse(). This function takes an array and performs what was described above. To note that this procedure is only rewarding if there is a great number of zeros. A downside to this approach is that in Matlab, the access to sparse matrixes is slower than when comparing with full matrixes.

### 9.4.2.2 Geometry Symmetries

LORs produced by a scanner are sometimes linked between themselves by geometrical symmetries. If exploited, these symmetries help to substantially reduce the amount of computation during the creation of the system matrix and computations related with it, since the memory allocated to the system matrix is reduced. Geometrical symmetries can be of different natures and exploited in different ways, mostly dependent on how the data is organized, and the reconstruction algorithm implemented.

For the easyPET system, given the data file used, the most computationally rewarding symmetry to implement was to divide a full revolution in four quadrants. By only creating a system matrix that gathers the LORs from the first quadrant, say 0º to 90º, it is possible to obtain the other three quadrants by means of multiple 90º rotations. This way, the system matrix becomes one quarter of a full, 0º to 360º, system matrix. This implementation is beneficial because the system matrix, by nature, needs to be allocated to a great amount of RAM. The greater the amount, the slower it becomes for a computer to access it. This way, one can greatly reduce the allocated space and optimize time consumption. The reason why smaller angular symmetries were not exploited is due to the perfect fitting of quadrangular pixels with the 90º symmetry interval. The increased implementation complexity and extra operations needed with, say, 45º symmetries, was undesirable.

## 9.5  ML-EM and Code Optimizations

As was introduced, the image reconstruction method designed for the easyPET system was based on the ML-EM algorithm. What follows is an attempt to expose the thought process behind its implementation, as well as the many optimizations that surround it.

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_i a_{ij}} \sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n} \qquad (9.7)$$

By observing the ML-EM algorithm Equation 9.7, one sees:

- three variables: $\lambda$ (reconstructed image), $a$ (system matrix), and $y$ (data);
- two indexes: $i$ $and$ $j$;
- four "short" calculations: $\frac{\lambda_j^n}{\sum_i a_{ij}}$, $y_i a_{ij}$, $\sum_{j'} a_{ij'} \lambda_{j'}^n$, and $\frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n}$;
- three "long" calculations: $\sum_i a_{ij}$, $\sum_{j'} a_{ij'} \lambda_{j'}^n$, and $\sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n}$.

The difference between "short" and "long" calculations resides on the usage of indexes. For "short" calculations, the indexes are fixed, and for "long" calculations it is needed a sweep throughout all indexes values, which is mathematically symbolized by the sigma sign $\sum$.

In terms of implementing an algorithm as complex as this, one can take two approaches. The first relies on making a highly descriptive handling of the problem, by computing each calculation separately, and performing each task independently. This is a symbolical approach, in which it translates the mathematical concept of solving such a problem by hand. Obviously, it would be inefficient to take this approach, as it does not take advantage of modern computational capabilities. The natural approach would be looking at the algorithm as sets of variables that can be calculated beforehand in order to take advantage of variable reusability and enhance task efficiency.

## 9.5.1 Building the Algorithm

Both back (top term) and forward (bottom term) operators are present in the following term of the ML-EM algorithm equation:

$$ImageNorm * \sum_i \frac{BackProjOper}{ForwardProjOper} = \frac{\lambda_j^n}{\sum_i a_{ij}} \sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n} \qquad (9.8)$$

### 9.5.1.1 Forward Projection Operation

The forward projection operator (bottom term) can be seen as independent of the index $j$ that is associated with the image iteration $\lambda_j^{n+1}$. This way, its implementation is set around a pair of nested for-loops that create an array called "sinotemp", with total length equivalent to the total number of unique activated LORs. Evidenced by the order of the summation signs, the nested for-loop sets the index $j'$ for a given index $i$ which, in return, is set by the outer for-loop.

```
for i = 1:numberOfBins

    for j = 1:numberOfPixels
        sinotemp(i) = sinotemp(i) + image(j)*sysmat(j, 1);
    end

end
```

**Figure 9.3:** Scripted code to exemplify the implementation of the Forward Projection operation.

### 9.5.1.2 Back Projection Operation

The back operator remains without nested summation signs. This being, it should only need one for-loop to be implemented. However, contrary to the previous operator, the back projection operator is not

independent of the index $j$. So, this operator still needs to be implemented by means of a pair of nested for-loops that create a new array variable called "imagecorr", with total length equal to the number of pixels in the image.

```
for j = 1:numberOfPixels

    for i = 1:numberOfBins
        imagecorr(j) = imagecorr(j) + counts(i)*sysmat(j, i);
    end

end
```

**Figure 9.4:** Scripted code to exemplify the implementation of the Back Projection operation.

### 9.5.1.3 Image Normalization

Much like the back operator, the image normalization operator (see Equation 9.8) is also dependent of the $j$ index, on top of also having a summation with index $i$. This operator can be easily implemented inside the nested for-loops depicted in Figure 9.5.

```
for j = 1:numberOfPixels

    for i = 1:numberOfBins
        imagecorr(j) = imagecorr(j) + counts(i)*sysmat(j, i);
        imagenorm(j) = imagenorm(j) + sysmat(j, i);
    end

    imagenorm(j) = image(j)/imagenorm(j);

end
```

**Figure 9.5:** Scripted code to exemplify the implementation of the Image normalization operation, inside the Back Projection operation.

### 9.5.1.4 Image Iteration

The final step is to iterate the image array, in this case portrayed by the array variable "image". The calculation of the forward operator is made previously to the back operator, so it can be introduced inside the back operator's for-loops. An extra for-loop is added to define the iteration exponent $n$.

```
for n = 1:numberOfIterations

    for i = 1:numberOfBins

        for j = 1:numberOfPixels
            sinotemp(i) = sinotemp(i) + image(j)*sysmat(j, 1);
        end

    end

    for j = 1:numberOfPixels

        for i = 1:numberOfBins
            imagecorr(j) = imagecorr(j) + counts(i)*sysmat(j, i);

            imagecorr(j) = imagecorr(j)/sinotemp(i);

            imagenorm(j) = imagenorm(j) + sysmat(j, i);
        end

        imagenorm(j) = image(j)/imagenorm(j);

        image(j) = imagenorm(j)*imagecorr(j);

    end

end
```

**Figure 9.6:** Scripted code to exemplify the implementation of the image reconstruction ML-EM algorithm. Note that an alteration is made inside the BackProjection Operation to accommodate the use of the operator that results from the Forward Projection, and, also, the image iteration.

## 9.5.2 Implementing Optimizations

Although the implementation explained above takes advantage of variable reusability, there is still much space for code optimization. Speeding up the algorithm is not necessary, yet desirable. One can perform this by implementing the following actions.

### 9.5.2.1 Not Computing Zeros

Since the FOV of the system is circular and the imaging space is quadrangular, roughly 22% of the imaging space is set to zero, this renders plenty of pixels set to zero in the cost function. On top of that, there are bins without any counts. Both situations produce zeros, which the algorithm takes time to compute without retrieving any result. Moreover, the zero values outside the FOV for either the image or system matrix, produce a division by zero and result in an error of "Not a number assignment" (Nan), in this case corresponding to the assignment of infinite values. To control any of these situations, the code was made robust by implementing a series of if-else blocks that would discard values which would not contribute to the image iteration.

### 9.5.2.2 Fast Indexing

As a way of speeding up the algorithm's computing, three matrix variables were created to centralize all the information needed for the algorithm, serving as variables for fast indexing. These variables were: "all_horizontal_sparse", "all_vertical_sparse", "uniq_pos".

"all_horizontal_sparse" – serves as a fast indexing variable for all the bins activated. It also stores the bin counts and serves as an indexer for "all_vertical_sparse";

"all_vertical_sparse" – as the previous, it serves as a fast indexing variable. In this case, aided by the previous, it aims to be used as a fast indexing to the entries in the system matrix, allowing a faster "look up" to the values in it;

"uniq_pos" – it provides a "look up" table for the bins that activate any given pixel. This meaning, if a pixel is activated, it goes and fetch the indexes of the produced LORs that might have activated it. It serves as a fast indexing matrix for the image and "all_horizontal_sparse".

These three variables have served a great purpose when reducing the time spent on image reconstruction. Although there has not been any observed bottleneck with these implementations, the only anticipated problem might relate with images that have highly disperse data, which would mean an overall increment on these matrixes sizes. This is yet to be observed.

Finally, the script for the ML-EM image reconstruction step was also set as a function. This enables an easier handling of the image reconstruction process and, later, enables the implementation of a user interface, as was discussed previously. Through the header of the ML-EM method function, the user controls the number of iterations to perform, the data file to be reconstructed and, as will be explained next, the image enhancement methods discussed in Chapter 5.2, image Gaussian filtering and the Median Root Prior.

## 9.6  Image Enhancement Implementations

Concerning the image enhancement methods discussed in Chapter 5.2, we will briefly introduce their implementation into the main body of the reconstruction algorithm. These methods comprise image Gaussian filtering and the implementation of a Median Root Prior (MRP).

### 9.6.1  Gaussian Filtering

Gaussian filtering was implemented via a native Matlab function called imgaussfilt(). This function is defaulted to spatial filtering and takes as input a 2D matrix. One can also change the standard deviation of the Gaussian, yet this was kept as the predefined value of $\sigma = 0.5$. The function was implemented before the keyword "end" of the iteration for-loop, after all the ML-EM operations described above, and controlled via an if-else block, so the user could select weather to use Gaussian filtering or not through the ML-EM function header.

### 9.6.2  Median Root Prior (MRP)

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_i a_{ij} + \beta \frac{(\lambda_j^n - M)}{M}} \sum_i \frac{y_i a_{ij}}{\sum_{j'} a_{ij'} \lambda_{j'}^n} \qquad (9.9)$$

The Media Root Prior is defined by the term underlined in yellow. It is inserted in the image normalization operation, and requires the calculation of a median value and a user defined $\beta$ paremeter. For the median value calculation, a neighborhood of the eight pixels surrounding a given pixel $\lambda_j$ was considered. For the definition of the parameter, user input is required in the form of a value between 0 and 1. Once the user inputs a $\beta$ value higher than 0, it acts with an if-else control, enabling the usage of the prior. The prior is added simultaneously to the image normalization operation.

## 9.7 Filtered Back Projection Implementation

For the Filtered Back Projection (FBP), a native Matlab function was used. This function, ifanbeam(), performs the Inverse Radon transform on a sinogram, taking into consideration the inputted geometry. As the name suggests, this function is predefined to perform the transform on the data originating from a fan-beam detector geometry. On top of that, it also allows to add a rotation to the fan-beam. This combination results in a transform fitted for the easyPET system's geometry.

For the input data, a sinogram like matrix was created. Yet, instead of using what was described in Chapter 3.3, a simpler approach was taken. This consisted in arranging a matrix with N by M entries, with N being the total number of top angle values, and M the total number of bot angle values. Each entry was then indexed the count value for each respective bin.

The ifanbeam() also takes other arguments as inputs. Mostly, since we are dealing with a method that uses filtering, it allows selecting the type of filter applied and parameters that characterize the filter used. For this work, and by definition, the ifanbeam() uses a "Ram-Lak" filter, or ramp filter, and "linear interpolation". It also lets the user control the outputted image and change the type of interpolation used for pixel value assignment during LOR back projection.

## 9.8 Image Quality Assessment and Algorithm Validation

An important step for every imaging system and its image reconstruction method is to assess whether they produce desirable results in a consistent manner. In the case of this work, this leads to algorithm validation and image quality analysis, the first being mainly performed by means of the second.

### 9.8.1 NEMA Rules

In order to have a standardized validation for imaging systems, the National Electrical Manufacturers Association (NEMA) has created a series of guidelines for medical imaging systems. The guidelines that concern this work are those meant for image quality assessment. Briefly, for small animals PET systems, it is required:

- Acquisition of data originating from a compact radioactive source;
- Reconstruction of acquired data using FBP;
- Spatial resolution for the point spread function (PSF) obtained must be measured in two directions, radial and tangential, being the PSF characterized by its Full-Width at Half-Maximum (FWHM) amplitude and Full-Width at Tenth-Maximum (FWTM) amplitude.

The guidelines also help define the necessary data and image characteristics:

- The radioactive source activity must be confined to no more than 3 mm in all directions and the radionuclide must be $Na^{22}$;
- Measurement must be acquired with the source located at the following radial distances from the center of the FOV: 5 mm, 10 mm, 15 mm, and 25 mm;
- The image pixel size must not be larger than one-fifth of the expected FWHM.

Provided these guidelines, one can see the importance of having implemented the FBP method. It is also important to state that the NEMA rules do not limit image reconstruction methods liable to be implemented for any given system. These rules aim only to standardize the systems validation, remaining up to the developers which other methods to implement. The importance of these rules will be discussed later.

### 9.8.2 Data for Reconstruction

The data used to test the easyPET system was obtained in two ways: through acquisition and simulation. Given the limitations imposed by the NEMA rules, Chapter 9.8.1, the data as to be accordingly. However, data acquisition is also limited by the precision one can position the radioactive source, as well as the availability of a given source. At the time of this work, acquired data had only been obtained with a point source, described before, in the positions: $0.56, 0.75, 6, 8.1, 10.3, 12.54,$ and 15 mm, from the center of the FOV. The latter being, simulated data was more convenient to obtain.

Simulation data was then obtained using GATE [50]. This platform enables simulation of PET data through Monte Carlo simulation, and by considering physical effects. Having the easyPET system designed in GATE, it was possible to position the source at the distances defined by the NEMA rules, and by those measured in the acquired data files. The reason such was made, was for comparison between acquired and simulated data. The results obtained would, eventually, justify the usage of simulated data when under the NEMA rules. Both the acquired and simulated data ranged from 4300 to 6300 counts.

### 9.8.3 Algorithm Validation

The validation of an algorithm is made mainly through the assessment of the quality of the images it produces. Event though, one can always evaluate the performance of a given algorithm by observing its computation times, and seeing if the implementation is robust and fit for reusability.

As was mentioned before, several functions were implemented in Matlab. The main purpose of this was to enable a user-friendly usage of the implemented method. This being, a Guided User Interface (GUI) was created in Matlab, allowing the user to input the reconstruction parameters according to the data used and the results to be obtained.

In terms of having a robust implementation, much like most software available, user interaction is of great importance. Through running the implemented method with several different data files and parameters, one can diagnose possible problems and solve them. The code is in constant update so as to address user feedback. Obviously, in such a linear software as this, the main arising problems are found during the implementation and test phase.

As for the implemented image quality assessment methods, two approaches were taken: signal-to-noise ratio (SNR) and, as demanded by the NEMA rules, spatial resolution.

### 9.8.4 Image quality assessment

For the image quality assessment, the analysis was made by calculating the image SNR and spatial resolution. Since the data used originates from circular 2D sources, a Matlab function was created in order to automate the full process of analyzing the several reconstructed image files. The scripted function loads each reconstructed image file and, on each one, draws an ellipse surrounding the source. With this ellipse are gathered the perimeter and the ellipse axes parameters, which help define the region of interest (ROI) of the source and the profiling lines.
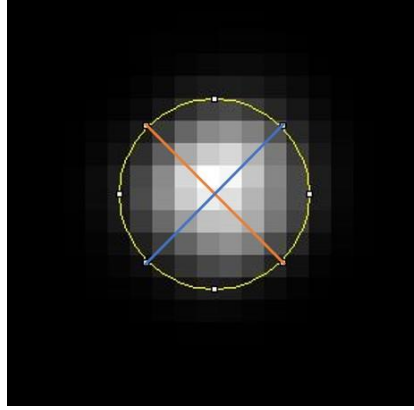
**Figure 9.7:** Schematic on how the ellipse ROI and the profile lines are defined.

### 9.8.4.1 Signal to Noise Ratio (SNR)

Using the ellipse perimeter, a regions of interest (ROI) was drawn, this comprising the source. With this ROIs is possible to calculate the needed quantities: the average signal value from the source and its standard deviation. This was made, contrary to using the background standard deviation, since the background yielded little to no signal, which would render an infinite SNR.

$$SNR = \frac{\mu_{source}}{\sigma_{source}} \qquad (9.10)$$

### 9.8.4.2 Spatial Resolution

Through the ellipse's axes parameters, it is possible to automate the drawing of two profile lines on top of the source, using the native Matlab function improfile(). After having the distance and respective intensity values, it is possible to fit a gaussian function to this data using the native Matlab function fit(). With the parameters of the gaussian curve we are able to calculate the FWHM and FWTM. The FWHM is calculated through:

$$FWHM = \sigma 2\sqrt{2ln(2)} \qquad (9.11)$$

And the FWTM:

$$FWTM = \sigma 2\sqrt{2\,ln(10)} \qquad (9.12)$$

Where, in both cases, $\sigma$ is the standard deviation of the obtained gaussian curve. This process was made in two directions, since it is required by the NEMA rules. The directions correspond to performing a profile along the x, and y axis of the ellipse.

# 10 Results

What follows are the results obtained for the described methods. First, an overview of the GUI implemented for the didactic system is made. After, the results obtained through image analysis are exposed. These results concern the determination of the ML-EM algorithm's iteration convergence; the testing of the system's data under NEMA rules, whilst comparing with the implemented iterative algorithm; finally, the implemented data regularization methods are also tested.

## 10.1 GUI

As was mentioned earlier, one of the goals of this work was to produce a user interface suitable for the didactic easyPET system. By substituting the scripted code and function headers, with buttons and input boxes, one enables a better user interaction and learning curve, with this system and with PET technique itself. This was made by taking advantage of Matlab's Graphical User Interface Development Environment (GUIDE), which allows creating Guided User Interfaces (GUI), effortlessly. The initial implementation was already revamped, so the GUI here presented is its current state. However, since the system's software is still in development and in constant change, the GUI serves mostly as a sketch for this idea and as a guideline for a more definite approach.

In Figure 10.1, we can see four images displayed, both obtained from the same data file: (1) is obtained via image reconstruction using the implemented ML-EM algorithm; (2) is a simple back projection image; (3) corresponds to the histogram of the data used to create the image (4) using the FBP method.

### 10.1.1 User Interaction

When the user clicks the "Load data" button, it is prompted to select the input data file. This action will automatically create images (2, 3, and 4) and fill out the "Parameters" section with the metadata available within the data file. At this step, the user can already observe changes produced in the FBP image when changing the filter used and its scaling, as well as changing the type of interpolation used.

After, the user can create the system matrix that corresponds to the data file, and with the desired pixels, defined in the "Start Reconstruction" section. Due to differences between acquired and simulated data files, the user is also required to select whether the inputted data file has resulted from an acquisition or simulation. After pushing the "Create System matrix" button, the user must wait for the system matrix to be concluded, a process that takes time, depending on the number of bins and pixels. Once the system matrix has concluded, the user can load it, or an alternative existing one, through the "Load sysmat" button.

To commence the image reconstruction process using the implemented ML-EM algorithm, the user is required to define the number of iterations to produce, and the amount of data he wants to reconstruct, defined by the number of turns, always an integer, equal or higher than one. It is also possible to select whether to perform image reconstruction with image quality enhancement methods (Non, Gaussian Filtering, or Median Root Prior – "Method" section). After checking if all the parameters are correct, the user can push the "Start" button and initiate the iterative image reconstruction method, a step that takes as much time as the number of pixels, bins, and counts, allow. Once this step is concluded, the user can input the resulting reconstructed images file and slide through the images produced at each iteration.
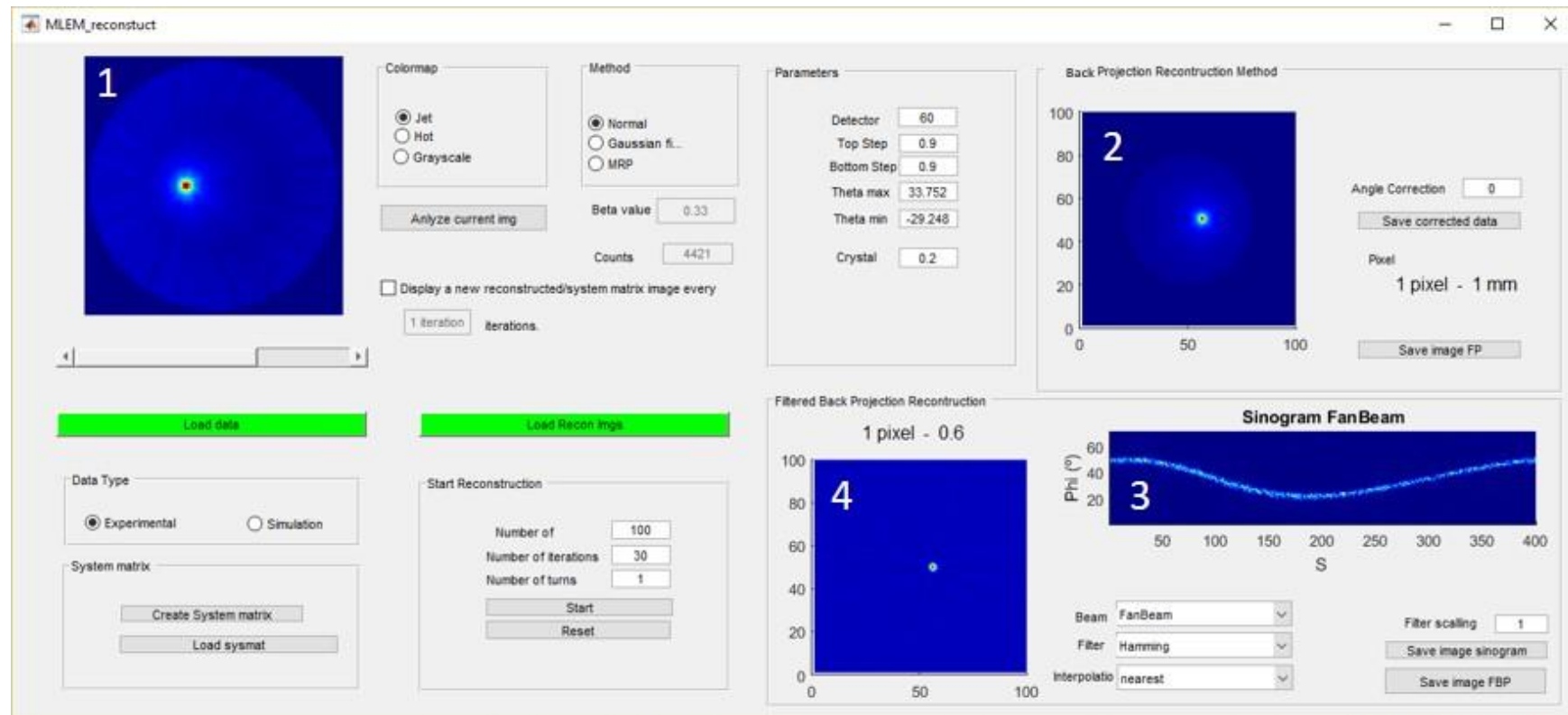
**Figure 10.1:** Set up of the GUI with data inputted as example. Images acquired with: (1) ML-EM algorithm; (2) Simple back projection; (3) sinogram representation; (4) Filtered Back Projection.

Finally, the user can change the color map used to see all images, and save each of the four presented images, individually. Additionally, the user can also push the "Analyze current img" button, which will temporarily save the current ML-EM image and open ImageJ with the saved image [51]. This software allows the user to perform several image analyses, including spatial resolution. The user can, at any time, push the "Reset" button and start over.

## 10.1.2 Advanced GUI

Since the work developed for this thesis is included in the easyPET project, which has seen its patent for the didactic system sold, and since a spin-off has been created around the technology developed, another user interface has been developed within this spin-off. This time, it was taken into consideration that by creating a software using freeware, one would avoid Matlab's fees, so that the availability of the easyPET software is greater. However, until this moment, this user interface only enables the user to control data acquisition, perform spectral analysis, observe live general information, and perform live image reconstruction using only back projection. Figure 10.2 represents the user interface created at RI.TE. Although it was not produced in the scope of this work, it remains as an example for a future and more definite version of the user interface for the easyPET system, both the didactic and pre-clinical versions.
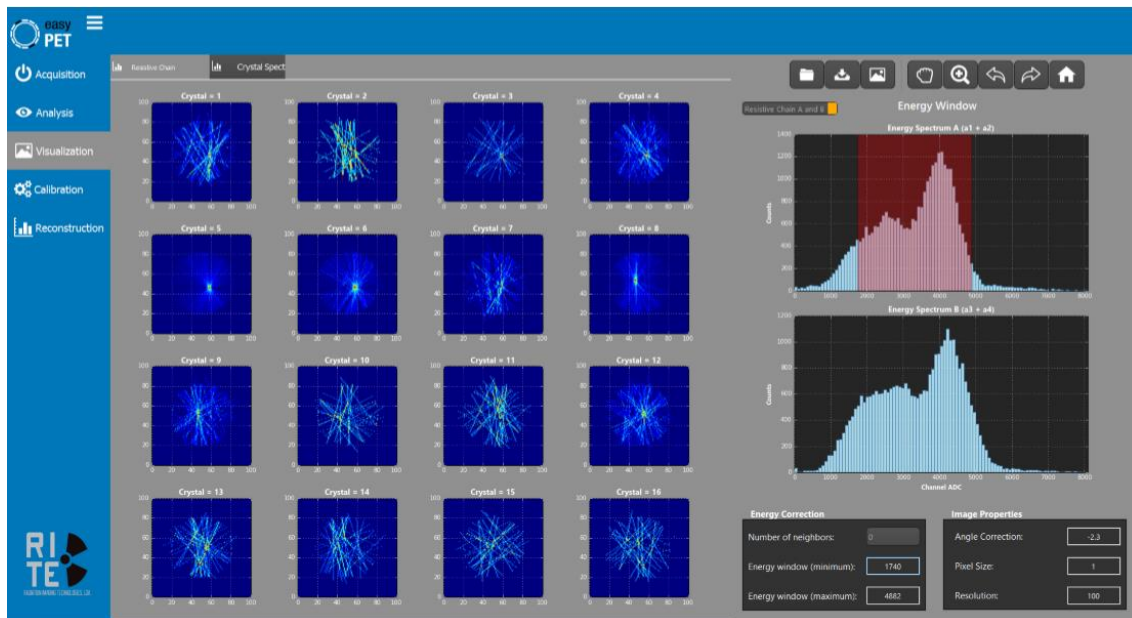


**Figure 10.2:** Current user interface developed. This interface allows control over the system acquisition, overview over the acquisition parameters, image statistics, and image back projection. This image was produced at RI.TE for the 1st Workshop on the Development of easyPET Technologies.

# 10.2 Image Analysis Results

As was mentioned in Chapter 9.8, spatial resolution at FWHM and FWTM, and for two different directions, was measured. The following results were obtained in order to:

1. See at what iteration the implemented algorithm would converge;
2. Observe any eventual differences between simulated and acquired data;
3. Test the data produced by the easyPET system, according to the NEMA rules defined in Chapter 9.8.1;
4. Compare the performances of the FBP method with the implemented native ML-EM image reconstruction algorithm;

5. Compare the performances of the native ML-EM algorithm *versus* the data regularization methods implemented, inter-iteration gaussian filtering and MRP.

Moreover, image analysis was also performed on images resulting from the use of data regularization methods. In this step, signal-to-noise ratio was also measured. In this last step, the comparison was made between images obtained through the native ML-EM algorithm, FBP, and the ML-EM algorithm using inter-iterations gaussian filtering and MRP using a $\beta$ parameter values of: $0.33, 0.66,$ and $1$.

Multiple acquisition and simulation files were obtained and, as we will see, the same spatial resolution behavior was observed between these files. This being, the results presented here on after correspond to the best produced results. Moreover, the working group estimated a spatial resolution between 1 and 1.5 mm, so that the pixel size was set to be around 0.25x0.25 mm$^2$, which led to the decision of using 230 x 230 pixels to describe the 57.7 x 57.7 mm$^2$ image space. The pixel size is justified in the NEMA rules (Chapter 9.8.1).

## 10.2.1 ML-EM convergence

Firstly, in order to determine the number of iterations needed to achieve convergence using the ML-EM algorithm, a series of runs with an excessive number of iterations and different data files, both simulated and acquired, were performed. Measuring the spatial resolution on the produced images enabled to best visualize the algorithm's convergence obtained. Figure 10.3 is an example for the behavior obtained for all data files, and it was admitted the algorithm reaches convergence at the 10th iteration. This assumption on the results was made by visually observing the stagnate evolution of both, FWHM and FWTM values in the several data files tested. This value was set as a reference, and its importance lies in not wasting time with excessive amounts of iterations as well as not incrementing image noise levels.
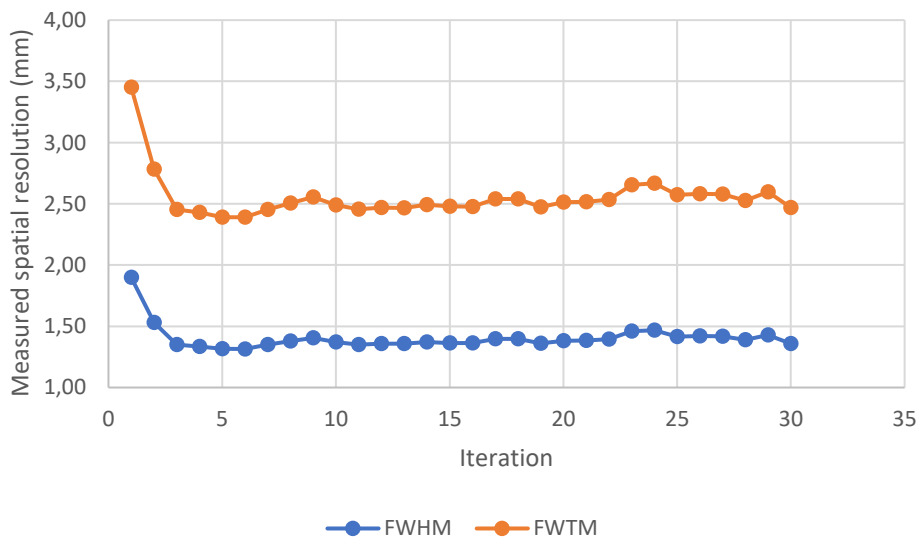


**Figure 10.3:** Graph illustrating the results obtained for FWHM and FWTM of the ML-EM algorithm, showing the algorithm's convergence around the 10th iteration.

## 10.2.2 Acquired *versus* Simulated Data

It is important to understand if there are any major differences between data acquired and simulated, so as to observe any faulty behavior in the physical easyPET system. As producing acquisitions requires more effort than producing simulations, already available acquisition data files with the source at specific distances, 0.56, 0.75, 6, 8.1, 10.3, 12.54, and 15 mm, from the center of the FOV, were matched with later simulated data at the same distances. Figure 10.4 shows the spatial resolution measured in the

x direction at FWHM, for the data files obtained in the aforementioned conditions, and for the FBP and ML-EM algorithm. The results obtained revealed that the same resolution levels can be achieved both data, using FBP or the ML-EM algorithm. Although, as can be seen in Figure 10.4, acquired data has produced results with larger variability. Moreover, it also produced resolution measurements, for most considered distances, significantly higher than those of simulated data. By taking into consideration the variability of the results obtained and the fact that the same resolution was measured at some distances. Allied with the significant difference in the number of coincidences detected for each data file type, depicted in Table 2.1, and the knowledge on PET data limitations, it is more than plausible to assume that the difference observed in both acquired and simulated data, comes from the statistics obtained. By having low numbers of coincidence registrations, each detected LOR weighs more in the final reconstructed image. This meaning, having, roughly, a difference of 1000 coincidences registrations between files, and with such low total coincidence numbers, if we assume the ratio of resolution-degrading LORs to be constant in simulated and acquired data, we were to expect a reduced resolution for simulated data. However, when considering that any unknown fault in the physical system is not implemented when producing simulated data, it is also plausible to say that the physical scanner presents some liabilities. To study this possibility, further testing is being conducted, which will be mentioned later. Due to this observed discrepancy in results, and since simulation enables positioning the source with greater precision, easyPET simulated data was tested under the NEMA rules, even though the NEMA rules specify measurements must be performed on acquired data.
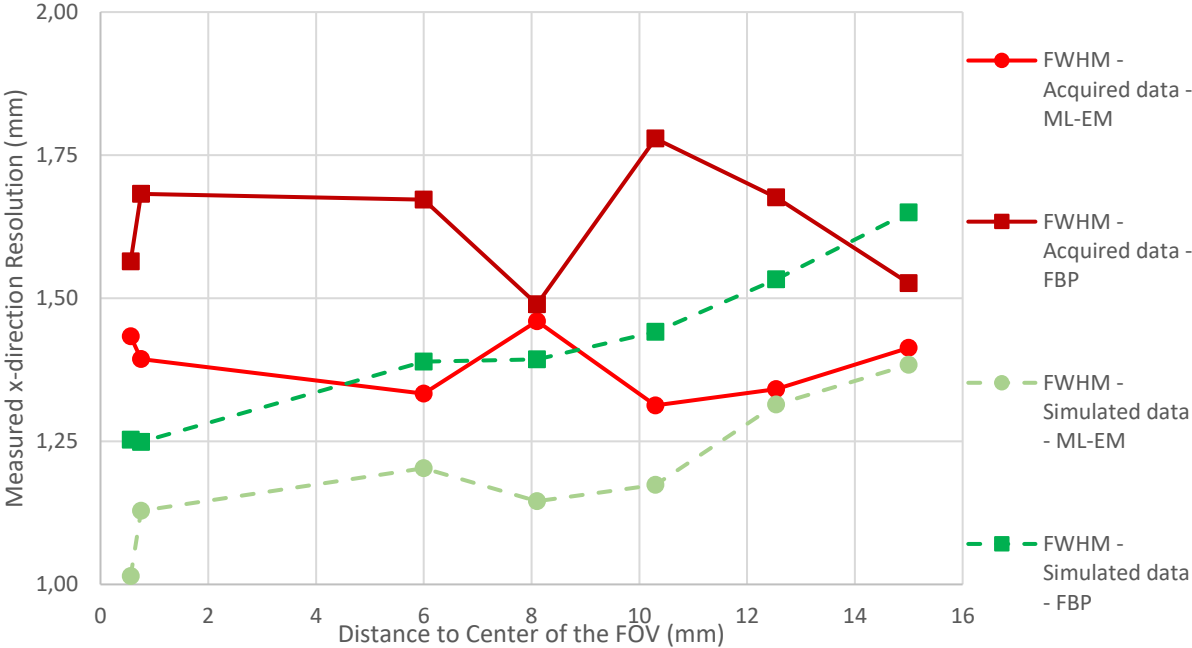


**Figure 10.4:** Graph illustrating the results obtained at FWHM. with the ML-EM algorithm and the FBP, using acquired and simulated data, at: 0.56, 0.75, 6, 8.1, 10.3, 12.54, and 15 mm from the center of the FOV.

**Table 10.1:** Number of coincidences detected for each simulated and acquired data file, at the specified distance from the center of the FOV.

| DISTANCE TO THE CENTER OF THE FOV (mm) | NUMBER OF COINCIDENCES DETECTED | |
|---|---|---|
| | Acquisition data | Simulated data |
| 0.56 | 4332 | 5702 |
| 0.75 | 4450 | 5723 |
| 6 | 4421 | 5687 |
| 8.1 | 4590 | 5856 |
| 10.3 | 4645 | 5811 |
| 12.54 | 4654 | 5891 |
| 15 | 4939 | 5901 |

## 10.2.3 NEMA Rules *versus* ML-EM Algorithm

Having previously determined that the ML-EM algorithm converges around its 10th iteration, and that simulation data was more suitable to test under the NEMA rules, the following results concern to reconstructing simulation data files under the NEMA rules explained in Chapter 9.8.1, and comparing it to the results obtained using the ML-EM algorithm, using the same data file.
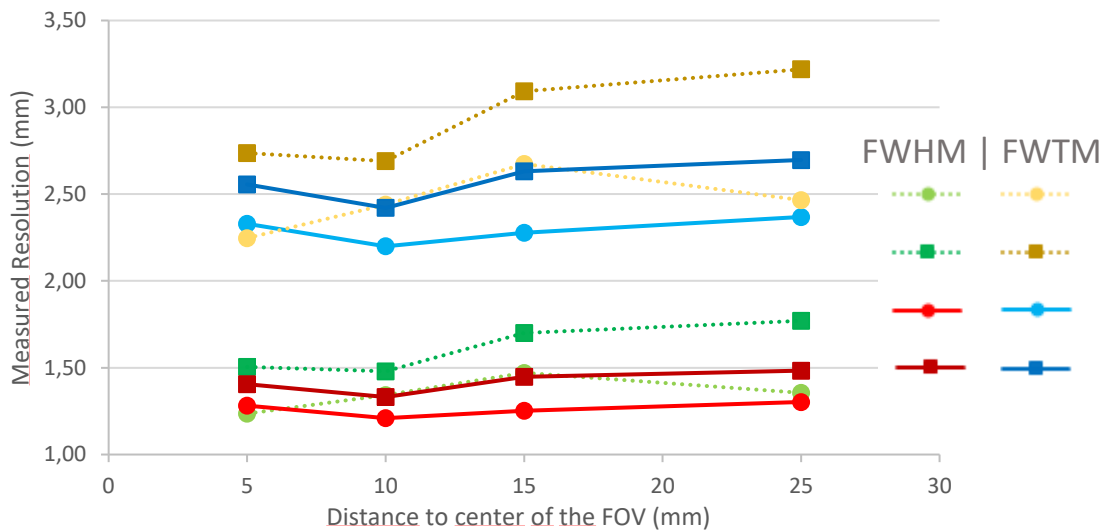


**Figure 10.5:** Graph illustrating the results obtained for a point source at: 5, 10, 15, and 25 mm from the center of the FOV. The FWHM and FWTM were measured in the x direction (dashed lines) and y direction (full lines), for the ML-EM algorithm (circle points) and FBP (square points).

The results shown in Figure 10.5 reveal that the ML-EM implementation (circle points) has achieved better and more uniform results than when using FBP (square points). By producing similar resolution measurements in the x (dashed lines) and y (full lines) directions, it is possible to say the iterative algorithm performs better in reconstructing the source's circular form, not contributing heterogeneously for both directions. This effect can be seen more clearly when observing the FWTM for both reconstruction methods (Figure 10.9), with the FBP producing a difference between the x and y directions of 0.5 mm or higher. Moreover, the higher FWTM values obtained with the FBP reveal that this method reconstructs the source's activity more dispersedly, resulting in a wider gaussian fit when performing the source's profile, corresponding to a higher ratio between the FWTM and FWHM values.

However, we see that at 10 and 15 mm from the center of the FOV, the difference between x and y directions resolution using the ML-EM method increased, whilst with FBP has not. Knowing that

positron annihilation produces a pair of photons emitted in a random direction, and that any PET data file has a limited number of counts, the reduced statistics associated with PET data is liable to produce biased results. In this case, it appears, when comparing to the remainder, that the simulation data files corresponding to these distances favor one direction over the other. In the case of FBP, it seems this effect is greater, the further away from the center of the FOV the source is. Allied with the filter convolution, which performs smoothing on the sources contour and might, contradict the effect observed with the ML-EM algorithm, at least to some extent.

## 10.2.4 Data Regularization Methods Comparison

After validating the ML-EM algorithm and comparing it with the FBP reconstruction under the NEMA rules, a comparison was made between reconstructing images using the iterative algorithm, whilst applying inter-iterations Gaussian filtering or the MRP. The MRP $\beta$ parameter was set to $0.33, 0.66$, and 1. The comparison was made by analyzing the produced FWHM and FWTM in two different directions, as described by the NEMA rules, and by calculating the SNR for the source. The objective of these comparisons is to assess and quantify what can be seen at each reconstructed image (e.g. Figure 10.6).
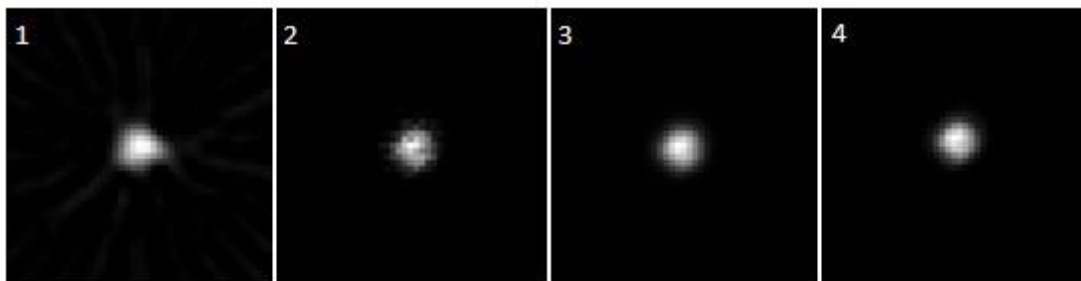


**Figure 10.6:** Example of how the point sources were reconstructed using: (1) FBP; (2) native ML-EM algorithm; (3) ML-EM algorithm with inter-iteration gaussian filtering; (4) ML-EM algorithm with MRP $\beta = 0.33$. Note that very little difference can be seen between (3) and (4), mainly due to the size of the source. (1) is clearly less round than the remainder and it is possible to see some LORs projected. (2) source appears less smooth.

### 10.2.4.1 Source Resolution

The following graph concerns to the comparison between the FWTM values obtained for the ML-EM algorithm without using quality enhancement methods, using inter-iterations Gaussian filtering, and using the MRP with $\beta = 0.66$. Only the FWTM is shown so as not to crowd the graph and since both the FWHM and FWTM have produced similar results. The same applies to the three $\beta$ values studied. As all produced similar results, it was chosen the one that produced the more evident, i.e. the one with lower resolution values throughout all distances to the center of the FOV. The FBP is not portrayed as it was already established that it produces higher resolution values than the iterative method, and it is expected that the implemented data regularization methods improve the algorithm results.

**Figure 10.7:** Graph illustrating the measured FWTM resolution values in x and y directions for the native ML-EM algorithm, ML-EM with inter-iterations gaussian filtering, and ML-EM algorithm with MRP and $\beta = 0.66$.

From Figure 10.7, it is possible to observe an improvement in the FWTM values, and in both directions, when using the implemented quality methods. However, we see that when using Gaussian filtering, the FWTM value in the y direction increases, comparing to the native ML-EM, whilst when using the MRP decreases. Moreover, both MRP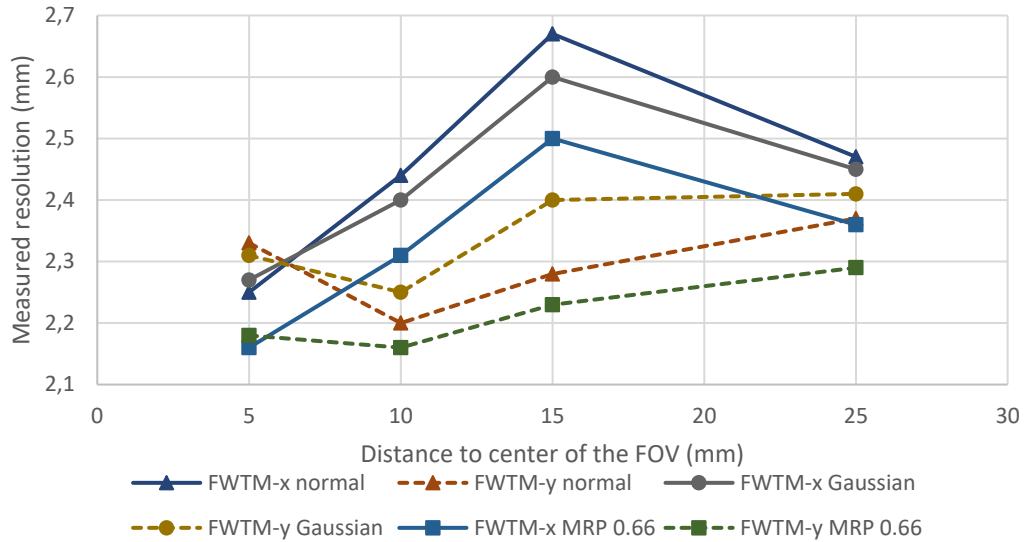 and gaussian filtering produce a reduced FWTM in the x direction. This ratio between the x and y directions (Figure 10.9) can be explained by the smoothing created when using gaussian filtering. This smoothing will result in a balancing between the x-y directions, by eliminating some of the irregularities in the contour of the source. In such small sources, each pixel contributes immensely when calculating the profile of the source, so that gaussian filtering produces such results. On the other hand, although the MRP produces lower resolution values, it does not tackle contour irregularities as strongly. This reflects the main property of the prior in producing an unaltered root image. Naturally, the sizes of the gaussian filter spatial kernels and the considered MRP neighborhood also take a role in producing different results. For such a small source, the smaller MRP neighborhood fits better to its contour. Visually, when observing Figure 10.6, it is easily seen that the FBP produces worst spatial resolution. However, qualifying spatial resolution between the other methods is more difficult.

### 10.2.4.2 Source SNR
Next, the SNR values for the FBP reconstruction, and ML-EM, with and without the implemented quality methods, were measured. This was also made throughout different distances to the center of the FOV: 5, 10, 15, and 25 mm. In this case, the $\beta$ values were set to: $0.33, 0.66,$ and $1$.
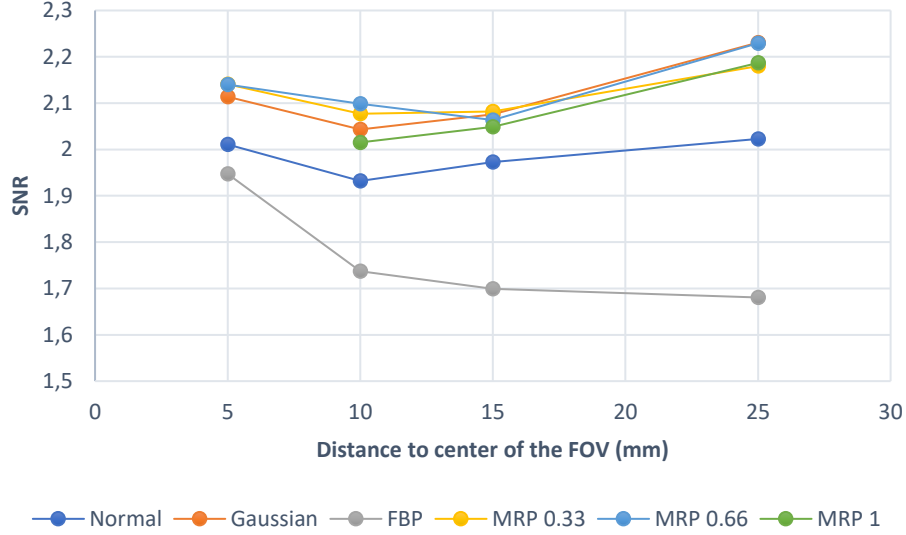
**Figure 10.8:** Graph illustrating the source SNR measured at different distances from the center of the FOV using the native ML-EM algorithm, the ML-EM algorithm using inter-iteration gaussian filtering, FBP, and ML-EM algorithm using MRP with $\beta = 0.33; 0.66; 1$.

In Figure 10.8, we see that the implemented image quality methods have yielded better SNR values than those produced by FBP and the native ML-EM implementations. The image quality methods have the property to reduce the presence of noise. As was seen in Figure 10.7, gaussian filtering produced a level of smoothing that contributed to a smaller disparity between the x-y directions resolution. In this case, the smoothing can be observed within the image source, producing a reduced standard deviation, and resulting in a higher SNR value. The same happens with the tested MRP values. However, with $\beta = 1$ at 5 mm from the FOV's center, the source's activity was reduced to a couple of pixels, hence not producing any SNR or resolution values. Since the remainder spatial positions have yielded positive results, this leads to the assumption that this $\beta$ value is not suitable in some situations. What can be seen from the MRP term (Equation 10.1) is that a low median value for a given high intensity pixel will lead to a prior with higher value. Eventually, this results in a lower reconstructed pixel value.

$$MRP = \beta \frac{\lambda_j - M}{M} \qquad ( 10.1 )$$

Overall, FBP reconstruction as yielded poor results when comparing with the remaining. It has already been noted that FBP yielded worst resolution results and, in this case, it can be seen it also produces noisier reconstructed source's, even though it takes advantage of filter convolution. It is also seen that the FBP performance varies with the distance to the center of the FOV, contrary to the remainder.

## 10.3 Final Discussion

In this work, the ML-EM image reconstruction algorithm was considered, and it yielded promising results. This algorithm was chosen based on a variety of factors. As it was mentioned before, with current technological advances and related increased computational power, iterative reconstruction algorithms have become widely used. On this notice, two major branches exist: algebraic and statistical algorithms. The main reason in choosing statistical algorithms relies on these taking into account the noise on the measured data. When dealing with a technique that produces small statistics, noise presence becomes very important to take into account. Algorithms such as Algebraic Reconstruction Technique (ART) are suitable for techniques with larger data and, consequently, increased computations, like CT. Amid iterative algorithms, the ML-EM has been widely used since it was first developed, and a good example

is the appearance of the OS-EM. The later was not taken onto consideration, over the ML-EM, as the easyPET is still in its initial phase, and data files still comprise small statistics. However, in the future, there might be room for this implementation, since the OS-EM produces good results and it is fairly simple to implement it over the ML-EM algorithm.

In Chapter 4.2.1 it was discussed that the ML-EM algorithm always converges to a solution. Yet, in terms of noise presence, it deteriorates for higher iteration numbers [14]. This being, it is important to determine the iteration number at which the algorithm converges and before it increments noise presence. This is exemplified in Figure 10.3, where we can see the ML-EM algorithm as converged around the $10^{th}$ iteration. However, the convergence is closely related with the data statistics used, rather than the algorithm itself. In this case, since the data statistics are fairly similar throughout the different data files used, comprising the same approximate amount of counts and with a distribution representing a point source, the iteration of convergence is set the same for all data files. In the case of more complex data (e.g. small animal data), where the data is more disperse and the statistics is higher, this analysis should be performed again.

On using acquired and simulated data, the main differences between both rely on the representation of physical events, and the precision with which the systems are described. Acquired data results from an existing equipment, with flaws and precision errors. Moreover, physical boundaries between and in media are represented "as is". The same goes with the randomness of physical events occurring whilst acquiring data. On the other hand, simulated data results from an attempt to best describe a real acquisition setup. In a simulation environment, the user inputs the characteristics of the media with exact control over them, as well as describing the scanner with a precision not possible when creating and assembling a physical scanner. This way, as much as it is possible to simulate the randomness of physical nature, it is only possible within a certain limit. This is why we can always expect to see less variability within simulated data, and why it is so important to further test image reconstruction methods using acquired data. This being, allied with the results observed previously in this Chapter, 10.2.2, led to the 2D easyPET system being, currently, under testing over the discrepancies in image resolution created by a position shift created by the coupling of the system's circuit board and the detectors. This is being made at the i3N lab, University of Aveiro, and the procedure is to apply several shifts to the top angle values, after acquiring the data on a point source, and observe the resolution behavior. The results obtained have been promising, since we have observed that the lowest resolution has been obtained when applying a shift. This means that, indeed, the detectors are not well aligned with the circuit board, and that the obtained differences, seen in Figure 10.4, already have a known source, not only the reduced statistics.

The results presented in Figure 10.5, show the potential of this system, when comparing with the machines introduced before in Chapter 6.2, Table 10.2**Table *6.1***. However, further image analysis is necessary. The limitations imposed and inherent to PET data can only be truly analyzed by testing the image reconstruction algorithms using more complex objects. Even though the NEMA rules have been defined to standardize PET systems and image reconstruction, the results obtained and the extensive bibliography on the subject have proven that image quality can be greatly enhanced using different image reconstruction algorithms, as well as by applying data regularization methods (Figure 10.7).

**Table 10.2:** Summary of the systems presented in Chapter 6 and, in addition, the results obtained with the easyPET system. *Algorithm used for spatial resolution measurement following NEMA rules. **Native algorithm developed for the system.

| SYSTEM | SPATIAL RESOLUTION (MM) | IMAGE RECONSTRUCTION USED |
|---|---|---|
| **DIGIPET** | 0.7 | MLEM*/3D-OSEM** |
| **NANOSCAN PET/MRI** | 1.5 | FBP*/3D-OSEM** |
| **$\beta$-CUBE** | 0.780 | FBP*/3D-OSEM** |
| **ALBIRA PET/SPECT/CT** | 1.65 | FBP*/3D-OSEM** |
| **EASYPET** | 1.5 | FBP*/2D-MLEM** |

In Figure 10.9 we can see the ratio between the measured FWTM for x and y directions, and for all the image reconstruction methods tested. A value closer to 1 represents a less eccentric source, i.e. reconstructed closer to reality. We can see that FBP produces the highest eccentricity, and Gaussian filtering has obtained more constant results, even though the data regularization methods go hand to hand. Moreover, FBP clearly presents worst results when the source is located further from the center of the FOV, whereas the remainder only present discrepancies at the 10 and 15 mm position. In this imaging technique, given the randomness of physic events, and the reduced statistics, it is possible to produce results that lack the characteristics of the real object. In the case of this point source study, it is important to notice that the smallest data contribution can severely change the resulting reconstructed source. Moreover, since the easyPET system does not have static detectors, their movement plus the randomness obtained from all events needed to produce PET data, might produce data files that do not reflect the nature of the object.
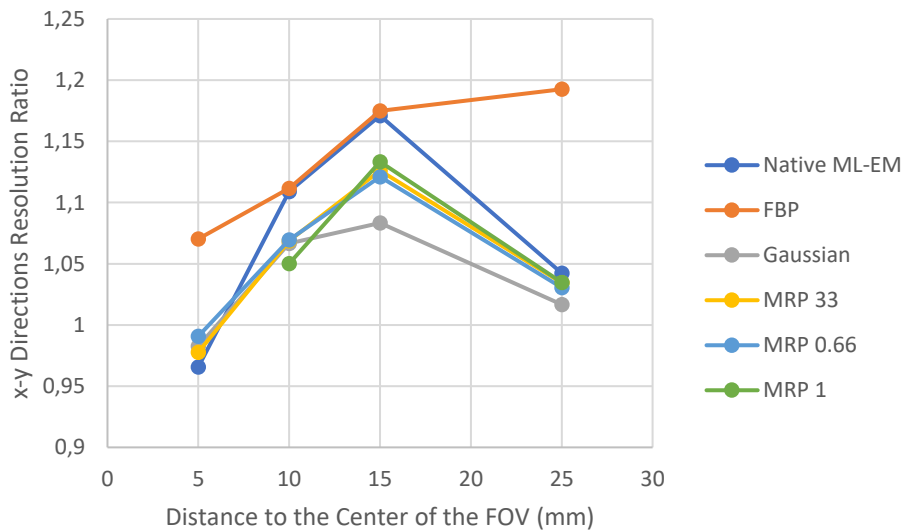


**Figure 10.9:** Graph showing the x-y directions resolution ratio for the measured FWTM and for all tested methods.

Regarding the implementations made in Matlab, it has been shown that the GUI has revealed to be a great tool for an enhanced user experience. Regarding the implementation of the image reconstruction algorithm and its optimizations, what was obtained is a robust implementation, performing well and yielding fast results. In terms of building the system matrix, using geometry symmetries as produced improved results with its reduction in time spent per reconstruction. As for the algorithm itself, the optimizations relating to fast indexing have, also, been rewarding.

# 11 Conclusion

In this thesis, it was presented the background that shows how PET technique works, and its importance, not only in clinical practice, but also in pre-clinical studies, and why there should exist more educational tools relating to its study. An important part of PET, much like any other imaging technique, is the need for image reconstruction. In the case of PET imaging, given the inherent characteristics of radioactivity and its detection, the nature of PET data is highly associated with noise presence. This being, the resulting reconstructed images will also carry noise. However, using specific image reconstruction methods, and by implementing data regularization methods, it is possible to reduce the effects created by the inherent noise. In the case of this work, the ML-EM algorithm was chosen. This iterative algorithm encompasses the Poisson nature of PET data, making it a statistic algorithm. This allows to make a series of consecutive approaches to a solution that better fits the data.

The motivation to this work was to produce an image reconstruction algorithm for the easyPET system. Given that the above was defined through the study of the existent bibliography, it was needed to fit an implementation of the ML-EM algorithm to the characteristics of the easyPET scanner. The later, being a highly innovative system, has required the study of its complex geometry, as well as the data it produces. The two opposing detector arrays, performing fan-like movements whilst rotating, posed as a challenge to implement with the ML-EM algorithm. Even more, it was uncertain whether it would produce data liable to have any potential in terms of future applications. However, since one of its designs was to be operated as a didactic tool, not required to achieved high quality results, it was thrilling to implement the image reconstruction algorithm for this, as well as to implement a user interface that served as a bridge between image reconstruction and the intended user.

Given the above, an implementation of the ML-EM image reconstruction algorithm was made for the 2D easyPET system. Since the ML-EM algorithm requires many computations, geometry symmetries of the easyPET system were exploited. Moreover, several code optimizations were performed. These two approaches have yielded significant results in reducing the initial computation type for the image reconstruction step. After the implementation of the algorithm, a user interface was created. This interface would, effortlessly, allow the user to interact with the image reconstruction process around the didactic system. Moreover, the user is enabled to apply the regularization methods implemented, as well as visually comparing the iterative reconstruction results with Filtered Back Projection.

In terms of the image quality produced by the system and the iterative algorithm itself, in this work the NEMA rules were presented. This set of rules allow a standardize evaluation to imaging systems and, in this case, also served as a standardized comparison between the ML-EM algorithm and the FBP method, when using the easyPET system. In terms of what was discussed in Chapter 10, the 2D easyPET system reveals much potential. In addition, the algorithm implemented in the scope of this project, has yielded better results than what was observed with the NEMA rules and FBP. As it was discussed earlier, further testing is needed on the easyPET acquired data to ensure the least amount of differences between acquired and simulated data.

Finally, the results obtained have shown that, as it stands, the didactic easyPET system fulfills the needs as an educational tool. The implementation of a user interface was a major factor in enhancing user experience with this system and PET technique. However, since the implemented scripts are still in constant change, the implemented GUI serves only as a sketch. Moreover, the results obtained for the reconstructed images reflect the potential of the pre-clinical system. Through further testing and implementations, it will be possible to either validate, or not, the commercial competitiveness of the easyPET system.

# 12 Future Work

With a multitude of PET systems available in the market, the easyPET arrives with a nouvelle approach to the common geometry, ensuring a cost reduction. This way, the easyPET project comes with the goal of tackling two issues: the lack of educational PET systems, and the ever increasing costs for pre-clinical scanners.

As it stands, the didactic approach has reach its final step, with this system's patent, comprising only a pair of detectors to produce 2D images, being already been sold. In this thesis, the implementation of the ML-EM image reconstruction algorithm for the 2D easyPET was presented, and the GUI implementation also left the guidelines for a future user interface implementation. What remains to do concerns to making the software for the didactic scanner more available, and easy to use. For this, the work performed in this thesis will continue, so as to implement the ML-EM algorithm, and the other parts of this work, in a freeware coding language. From the point of view of RI.TE, a user interface is in development, and, later on, it will be possible to attach the image reconstruction step to it. The final stage, which will enhance the usability of the didactic system, is the creation of a simulation package for GATE [50], so that the availability of radioactive sources is not an obstacle, and people can make the best of the didactic easyPET system.

The pre-clinical easyPET scanner is at a more intermediate state. With regards to a 2D functional mode, it is in the same stage as its didactic version. However, the pre-clinical scanner goals are to compete against current pre-clinical systems. For this, new hardware implementations are under study. The current system already has two column arrays with 16 detectors each, and a scaled up version has been idealized, having two modules of 50x5 LYSO scintillators. Moreover, since the detector's size is a major limitation of the system's resolution, $1.5x1.5x20$ mm$^3$ crystals are currently under study. Regarding the software, the pre-clinical system is at the same stage as its didactic version. However, it still requires a fully implemented 3D image reconstruction algorithm. As for this moment, the ML-EM algorithm has been implemented for multislice reconstruction, and data rebinning is under study. The fully 3D implementation is almost concluded. The step after relates to implement all the related software in freeware.

From this work, and the interest it has gathered since its appearance, we can see that it has a future filled with potential. What remains is to keep on working on it to ensure it happens, and, hopefully, the easyPET concept will prevail.

# 13 References

[1] "American Cancer Society | Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin." [Online]. Available: https://www.cancer.org/. [Accessed: 26-Jun-2017].

[2] N. R. Carlson, *Physiology of behavior: Methods and Strategies of Research*, 11th ed. Allyn & Bacon, 1986.

[3] R. A. Serway, C. J. Moses, and C. A. Moyer, *Modern Physics*. Cengage Learning, 2004.

[4] J. Langner, "Event-driven motion compensation in positron emission tomography: Development of a clinically applicable method," 2008.

[5] "Photomultiplier," *Wikipedia*. 18-Jul-2017.

[6] G. F. Knoll, *Radiation Detection and Measurement*. John Wiley & Sons, 2010.

[7] T. K. Lewellen, "Recent developments in PET detector technology," *Phys. Med. Biol.*, vol. 53, no. 17, pp. R287–R317, Sep. 2008.

[8] M. N. Wernick and J. N. Aarsvold, *Emission Tomography: The Fundamentals of PET and SPECT*. Academic Press, 2004.

[9] "Data Structure | UBC-TRIUMF Positron Emission Tomography." [Online]. Available: http://www.pet.ubc.ca/data-structure. [Accessed: 25-Jul-2017].

[10] D. Townsend and M. Defrise, "Image Reconstruction Methods in Positron Tomography. In Lectures given in the Academic Training Programme of CERN." 1993.

[11] S. R. Deans, *The Radon Transform and Some of Its Applications*. Courier Corporation, 2007.

[12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc. Ser. B Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[13] Y. Vardi, L. A. Shepp, and L. Kaufman, "A statistical model for positron emission tomography," *J. Am. Stat. Assoc.*, vol. 80, no. 389, pp. 8–20, 1985.

[14] S. Vandenberghe *et al.*, "Iterative reconstruction algorithms in nuclear medicine," *Comput. Med. Imaging Graph.*, vol. 25, no. 2, pp. 105–111, Mar. 2001.

[15] R. M. Lewitt and S. Matej, "Overview of methods for image reconstruction from projections in emission computed tomography," *Proc. IEEE*, vol. 91, no. 10, pp. 1588–1611, 2003.

[16] R. Fisher, "On the probable error of a coefficient of correlation deduced from a small sample," *Metron*, vol. 1, pp. 3–32, 1921.

[17] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc. Ser. B Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[19] S. K. Sengijpta, "Fundamentals of Statistical Signal Processing: Estimation Theory," *Technometrics*, vol. 37, no. 4, pp. 465–466, Nov. 1995.

[20] D. S. Lalush and N. W. Miles, *Iterative Image Reconstruction-Chapter 21*. 2004.

[21] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, Dec. 1994.

[22] J. Sanders and E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming, Portable Documents*. Addison-Wesley Professional, 2010.

[23] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Elad, and M. Israeli, "Fast and accurate Polar Fourier transform," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 2, pp. 145–167, Sep. 2006.

[24] H. Peng and C. S. Levin, "Recent Developments in PET Instrumentation," *Curr. Pharm. Biotechnol.*, vol. 11, no. 6, pp. 555–571, Sep. 2010.

[25] "PET/CT and Attenuation Correction Artifacts." [Online]. Available: http://www.people.vcu.edu/~mhcrosthwait/PETW/PETCTattenuationcorrectionandartifacts.html. [Accessed: 24-Jul-2017].

[26] S. Vandenberghe, E. Mikhaylova, E. D'Hoe, P. Mollet, and J. S. Karp, "Recent developments in time-of-flight PET," *EJNMMI Phys.*, vol. 3, no. 1, p. 3, Feb. 2016.

[27] "Introduction to PET Physics: The physical principles of PET." [Online]. Available: https://depts.washington.edu/nucmed/IRL/pet_intro/intro_src/section2.html. [Accessed: 24-Jul-2017].

[28]  M. E. Casey and E. J. Hoffman, "Quantitation in positron emission computed tomography: 7. A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration.," *J. Comput. Assist. Tomogr.*, vol. 10, no. 5, pp. 845–850, 1986.

[29]  D. L. Bailey, D. W. Townsend, P. E. Valk, and M. N. Maisey, *Positron emission tomography*. Springer, 2005.

[30]  C. Knoess *et al.*, "Evaluation of the Depth of Interaction (DOI) for the High Resolution Research Tomograph (HRRT)-a comparison between scanners with and without DOI," in *Nuclear Science Symposium Conference Record, 2002 IEEE*, 2002, vol. 3, pp. 1447–1451.

[31]  M. V. Green, H. G. Ostrow, J. Seidel, and M. G. Pomper, "Experimental Evaluation of Depth-of-Interaction Correction in a Small-Animal Positron Emission Tomography Scanner," *Mol. Imaging*, vol. 9, no. 6, pp. 311–318, Dec. 2010.

[32]  R. Gonzalez and R. Woods, *Digital Image Processing*. Addison-Wesley Publishing Company, 1992.

[33]  S. Alenius and U. Ruotsalainen, "Bayesian image reconstruction for emission tomography based on median root prior," *Eur. J. Nucl. Med.*, vol. 24, no. 3, pp. 258–265, Mar. 1997.

[34]  D. H. Yom and S. Ann, "Directed graph representation for root-signal set of median filters," *Proc. IEEE*, vol. 75, no. 11, pp. 1542–1544, Nov. 1987.

[35]  S. Alenius, U. Ruotsalainen, and J. Astola, "Using local median as the location of the prior distribution in iterative emission tomography image reconstruction," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 3097–3104, 1998.

[36]  "pet_camera_3d_view.png (1024×768)." [Online]. Available: http://www.imalone.co.uk/images/pet_camera_3d_view.png. [Accessed: 28-Jul-2017].

[37]  S. España, R. Marcinkowski, V. Keereman, S. Vandenberghe, and R. Van Holen, "DigiPET: sub-millimeter spatial resolution small-animal PET imaging using thin monolithic scintillators," *Phys. Med. Biol.*, vol. 59, no. 13, pp. 3405–3420, Jul. 2014.

[38]  P. Mollet *et al.*, "The β-CUBE, a high-end compact preclinical benchtop PET for total body imaging," *J. Nucl. Med.*, vol. 58, no. supplement 1, pp. 393–393, May 2017.

[39]  K. Nagy *et al.*, "Performance Evaluation of the Small-Animal nanoScan PET/MRI System," *J. Nucl. Med.*, vol. 54, no. 10, pp. 1825–1832, Oct. 2013.

[40]  F. Sánchez *et al.*, "Small animal PET scanner based on monolithic LYSO crystals: Performance evaluation," *Med. Phys.*, vol. 39, no. 2, pp. 643–653, Feb. 2012.

[41]  V. Arosio *et al.*, "The easyPET: a novel conceptfor an educational cost-effective positron emission 2D scanner."

[42]  V. Arosio *et al.*, "easyPET: a novel concept for an affordable tomographic system," *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.*, vol. 845, pp. 644–647, Feb. 2017.

[43]  "SP5700 / Nuclear & Modern Physics Kits | CAEN." [Online]. Available: http://www.caen.it/csite/CaenProd.jsp?idmod=1025&parent=61. [Accessed: 25-Jun-2017].

[44]  "RI-TE, Lda." [Online]. Available: http://www.ri-te.pt/index.html. [Accessed: 27-Jun-2017].

[45]  "i3N." [Online]. Available: http://www.i3n.org/. [Accessed: 25-Jun-2017].

[46]  "MathWorks - Makers of MATLAB and Simulink." [Online]. Available: https://www.mathworks.com/. [Accessed: 25-Jun-2017].

[47]  "STIR Home Page." [Online]. Available: http://stir.sourceforge.net/homepage.shtml. [Accessed: 25-Jun-2017].

[48]  M. Jacobson *et al.*, "Enhanced 3D PET OSEM reconstruction using inter-update Metz filtering," *Phys. Med. Biol.*, vol. 45, no. 8, p. 2417, 2000.

[49]  G. T. Herman, *Image Reconstruction from Projections*. 1979.

[50]  "Forewords | GATE." [Online]. Available: http://www.opengatecollaboration.org/. [Accessed: 27-Jun-2017].

[51]  "ImageJ." [Online]. Available: https://imagej.nih.gov/ij/. [Accessed: 27-Jun-2017].